VOL. XVI AUGUST, 1932 No. 4

The Journal of APPLIED PSYCHOLOGY

JAMES P. PORTER, EDITOR
Ohio University
WITH A BOARD OF CO-OPERATING EDITORS

TABLE OF CONTENTS

M. O. WILSON and J. H. HODGES. Predicting Success in the Engineering College 343
HAROLD E. BURTT and HAROLD V. GASKILL. Suggestibility and the Form of the Question 358
ROBERT H. THOULESS. Professor H. F. Adams and The Two-Factor Theory of Ability 374
H. T. MANUEL and LOIS S. HUGHES. The Intelligence and Drawing Ability of Young Mexican Children 382
DONALD G. PATERSON and MILES A. TINKER. Studies of Typographical Factors Influencing Speed of Reading: VIII. Space Between Lines or Leading 388
K. H. R. EDWARDS. The Rating of Temperamental Qualities 398
RICHARD LEDGERWOOD. Measurement of the Appeal of Performances in the Theater 405
HAROLD D. GRIFFIN. Constructing a Prediction Chart 406
EDNA WILLIS MCELWEE. The Primary Cylinder Board 413
FLORENCE BAKER and M. E. BROOM. Concerning One Criterion for the Choice of Primary Reading Tests 419
GRACE S. RODGERS. A Critical Study of the Grammatical Errors of Junior High School Pupils 421
LOTTA M. KARPELES. Porteus Maze as a Discriminative Measure in Delinquency 427
Notes and News 438
Book Reviews 440
New Books and Pamphlets 451

PUBLISHED BI-MONTHLY BY THE EDITOR
Ohio University, Athens, Ohio

Printed by
THE SCIENCE PRESS PRINTING COMPANY
LIME AND GREEN STS., LANCASTER, PA.

Entered as second-class matter at the post office at Athens, Ohio, under the act of March 3, 1879. Additional entry as second-class matter at the post office, Lancaster, Pa. Acceptance for mailing at special rate of postage provided for in section 1103, act of October 3, 1917. Authorized December 21, 1922.

COPYRIGHT 1932, BY JAMES P. PORTER

The Journal of Applied Psychology

CO-OPERATING EDITORS

W. F. Book, Indiana University, W. V. Bingham, Personnel Research Federation, New York; June E. Downey, University of Wyoming; Robert H. Gault, Northwestern University; William Healy, Baker Foundation, Juvenile Court, Boston; H. L. Hollingworth, Columbia University; J. E. W. Wallin, Atlantic University; Herbert S. Langfeld, Princeton University; James H. Leuba, Bryn Mawr College; Henry M. McComas, Princeton University; Walter D. Scott, Northwestern University; C. E. Seashore, University of Iowa; Daniel Starch, Boston, Mass.; Edward K. Strong, Jr., Leland Stanford Jr. University; L. M. Terman, Leland Stanford Jr. University; E. L. Thorndike, Columbia University; Joseph Jastrow, New School of Social Research, New York; G. M. Whipple, Danvers, Mass.; Helen Thompson Wooley, Teachers College, Columbia University.

The scope of The Journal of Applied Psychology will include the following: (a) The application of psychology to vocational activities, such as law, art, public speaking, industrial and commercial work, and problems of business appeals. (b) Studies of individual mentalities, such as types of character, special talents, genius, and individual differences, including the problems of mental diagnosis and vocational prognosis. (c) The influence of general environment conditions, such as climate, weather, humidity, temperature; also such conditions as nutrition, fatigue, etc. (d) The psychology of everyday activities, such as reading, writing, speaking, singing, playing games or playing musical instruments, sports, etc.

Communications, manuscripts, reviews, summaries, notices, books and off-prints for review should be addressed to Dr. James P. Porter, Ohio University, Athens, Ohio. Contributors are requested to follow the conventional method of citation employed by this Journal; in every case the year, volume and pages should be specified.

The Journal of Applied Psychology is published bi-monthly. The subscription price is $5.00 a year. Single numbers $1.00 to $2.00 depending upon the number of pages. Remittances and business communications should be addressed to the Editor, James P. Porter, Ohio University, Athens, Ohio.

PREDICTING SUCCESS IN THE ENGINEERING COLLEGE

M. O. WILSON AND J. H. HODGES
University of Oklahoma

This report presents the results of a study in the prediction of success in the Engineering College of the University of Oklahoma. The purpose of the study is to determine to what extent it is possible to predict at the end of the freshman year the success the students will have in pursuing their advanced engineering courses required for completing a degree. If the prediction can be made with a fair degree of accuracy, the results of such a study should be of some value in directing the work of students matriculated in the engineering college.

HISTORICAL

It will not be practical here to undertake a complete review of all the studies which have been made on the problem of predicting scholastic success; however, a few such reviews may not be out of place.

An early study which indicated an interest in predicting scholastic success was that of Lewis (6) in 1905. In a paper on formal discipline he gave the results of some observations he made concerning the relation of success in mathematics to subsequent success in law courses at Dartmouth. Without the use of any rigorous statistical technique he found that 50 percent of the best students in law were conspicuous for their poor showing in mathematics and that 42 percent of those poorest in law stood at the head of the group ranked according to ability in mathematics. He concluded that these results convince one that students able in mathematical reasoning are not even generally able in practical reasoning in law.

Later, however, Rietz (13) challenged the results of Lewis. He took the scholarship records of Dartmouth law classes from 1897 to 1901, including some of the same students studied by Lewis, and correlated mathematics with law, obtaining $r = .528$ ($N = 161$). So, his conclusion was somewhat the opposite of that of Lewis, namely, that as a result of this significant positive correlation it is possible to predict with a fair degree of accuracy success in law courses when achievement in mathematics is known.

This finding also agrees with that of Groves (2), who studied the relation of grades in freshman prescribed arts and science courses to those of law courses. The freshman courses and their coefficients of correlation with law are as follows: science (physics, chemistry, zoology, botany), .63; mathematics, .46; modern language, .38; English, .34; and government, .30 ($N = 107$). In these results it is seen that mathematics is second only to science in its power of predicting success in law.

Thurstone (15) has suggested that admission to college should be determined by the following criteria: high school credits, high school principal's estimate of capacity, college entrance examination, dean's interview with the applicant, and psychological tests. He further suggests that a method of critical scores is more controllable than the multiple correlation technique for establishing diagnostically significant criteria.

May (8) made a study of predicting success in academic subjects with the purpose, among other things, of determining which variable or combination of variables was the most reliable in such prediction. The prognostic measures selected were scores on the Miller Mental Ability Test and the Dartmouth Completion of Definition Test, the number of high school units offered, high school grades, and application as measured by the number of hours per week spent in study by each student. The coefficients of correlation of each of the independent variables with the criterion are: intelligence,

.60; high school units, .22; high school grades, .40; combined high school units and grades, .24; application, .32. For the combined predictive instrument, R = .84; when units were omitted, R = .838; when high school units and grades were omitted, R = .825; and in neither case was the error of estimate materially increased by the omission. The author concluded, contrarily to Thurstone, that high school grades and high school units may be omitted from the predictive instrument and emphasized the importance of intelligence and application. However, it would be necessary to secure information concerning application at about mid-term or later, as such information is not available at the time of matriculation. Otherwise, intelligence alone should be used.

Rector (12) studied the prediction of success in high school and used as his independent variables Army Alpha, scholarship ratings, and application ratings. These correlated respectively with the criterion, high school grades, as follows: Army Alpha, .28; scholarship ratings, .25; and application ratings, .25. The combination of these independent variables correlated with the criterion to the extent of R = .34 and it was concluded that they are not valid for predicting high school success.

Hull and Limp (5) constructed batteries of tests in order to predict individual aptitudes in shorthand, typewriting, English and algebra in high school. The batteries correlated with their respective criteria in the order listed above to the extent of .51, .61, .65 and .74 (N = .73). However, when these tests were used to measure the *difference* between a person's aptitude to do one thing and a second thing, the coefficients dropped to .17, .23, .31 and .42. This shows that to distinguish the various degrees of different aptitudes within a single individual is a different and more difficult problem than to distinguish amounts of a single aptitude in different individuals. The outlook is promising for such a test, however.

In a subsequent article Limp (7) attempted to determine not individual differences in aptitude for English and short-

hand but the difference between the aptitudes of a single individual for the two types of work. The battery of tests with their corresponding weights in the regression equation derived were best answers, +.490; sentence meaning, +.280; number series, +.321; spelling, +.092; and a non-designated test, +.171. The multiple coefficient of the combined device was R=.518. From these results the author was of the opinion that a regression equation can be constructed which will predict with a fair degree of accuracy the *difference* between aptitudes in an individual.

In a study by Nelson and Denny (10), a comparison was made of the relative predictive power of two thirty-minute tests, namely Thurstone's Group Test of Mental Ability, Form IV, and Terman's Group Test of Mental Ability. Extent of agreement between the tests was represented by $r = .66 \pm .02$. With scores on two different objective examinations in Psychology 1, the intelligence tests correlated as follows: Terman, $.64 \pm .06$, and $.77 \pm .02$; Thurstone, $.51 \pm .08$ and $.60 \pm .05$. When correlated with general scholarship of the first semester of the freshman year, the intelligence tests yielded the following coefficients: Terman, $.45 \pm .05$ and Thurstone, $.49 \pm .05$. The authors concluded that both tests are fairly satisfactory in predicting success in Psychology 1, with a decided advantage in favor of Terman. In predicting general success both tests are mediocre but compare favorably in this respect to tests requiring longer time for administering.

Pierson and Nettels (11) attempted the prediction of college success, using as a predictive instrument a battery consisting of high school grades, IQ (Terman Group Test), and teachers' ratings on character. These correlated respectively with the criterion as follows: .523, .429, and .410 (N=50). The multiple coefficient for this battery was R=.650.

A suggestion in making a prognostic battery has been made by Spearman (14), Holzinger (3), Asher (1), McDonough (9), and others, to the effect that tests should satisfy the Spearman tetrad criterion. A higher correlation between the

battery and the criterion can be obtained if the correlations between the battery items depend upon *g*, since it appears that school grades are in large measure dependent upon *g*.

METHOD OF PRESENT STUDY

The group. Beginning with the fall semester of 1923–1924 the Otis Advanced Intelligence Scale was administered annually for a number of years to the freshmen on entering the College of Engineering of Oklahoma University. One hundred twenty-five freshmen were selected between the years of 1923 and 1926, inclusive, for this study. However, on examining their scholarship records it was found that of this group eighteen were irregular in their freshman courses and were not comparable to the remainder. Therefore, the group finally used consisted of one hundred seven students who had taken the Otis test, had completed all of the prescribed freshman subjects and, with the exception of seven who had withdrawn during the junior or senior year, had completed the four-year course and received an engineering degree.

The criterion. Success in engineering courses is measured by the grades the students receive in these courses—these are the only measures available. An attempt is made to combine quality and quantity of work by use of honor points which are contributed to by the amount of work done in terms of hours and by the quality in terms of grades. The criterion or dependent variable is the average of honor points in all courses taken beyond the freshman year. The advanced work in the college is divided into the following schools or departments, each offering a degree: chemical engineering, civil engineering, electrical engineering, engineering physics, geological engineering, mechanical engineering, petroleum engineering, mining engineering, and architectural engineering. The results would have been more valuable if the predictions could have been made independently for each of the different departments. However, this would have increased the statistical labor about nine fold and would have correspondingly

decreased the reliability of the results, since the limited number of students for whom necessary data were available was too small to admit of such detailed analysis. The coefficient of reliability of the criterion is .560. This was obtained as follows: The grades in each of the engineering departments were arranged in the order in which the courses were taken. Then the first half were correlated with the second half. These coefficients ranged from .42 to .64. Finally these departmental coefficients were averaged, giving .56. At best this is a very crude method but seemed to be the only one available.

The basis of prediction. There are many factors which contribute to success in any college such as general ability, high school preparation, mental and physical health, efficiency of study habits, course load, application, personality and character traits, and the social environment of the student. Since the influence of most of these is not known the independent variables to be used for predicting success in the advanced engineering work are the Otis score and the average scholarship in certain freshman prescribed courses. This prescribed plan of work, which is common to all of the departments named above, is comprised of the following: introductory engineering, one semester; English, (composition), two semesters; mathematics (college algebra, trigonometry, and analytics), three semesters; mechanical drawing, two semesters; chemistry (general and qualitative), two semesters; wood and forge work, two semesters; and physical education (or military science), two semesters. The latter two were omitted since it was assumed that achievement in these courses would show less relation to subsequent achievement in engineering than would the others.

In Table 1 are shown the means, standard deviations and coefficients of reliability for the criterion and the variables used for prognosis. The honor points assigned arbitrarily to each letter grade per semester hour of credit are as follows: A = 7, A − = 6, B = 5, B − = 4, C = 3, C − = 2, D (condition) = 1,

TABLE 1

Variables tentatively used in predicting success in advanced engineering courses of 107 students in the University of Oklahoma

VARIABLE	MEAN	STANDARD DEVIATION	COEFFICIENT OF RELIABILITY
1. Advanced engineering average (criterion)	4.360	0.933	.560
2. Freshman engineering	4.482	1.428	.....
3. Freshman English	4.444	1.153	.506
4. Freshman mathematics	4.312	1.364	.569
5. Freshman mechanical drawing	4.816	1.184	.588
6. Freshman chemistry	4.322	1.268	.413
7. General freshman average	4.355	1.040	.....
8. Otis intelligence score	160.830	24.191	.967

F (failure) = 0. (Since these data were compiled the number of passing grades in the university has been reduced from six to four). No attempt was made to reduce grades in the various departments to a comparable basis. Neither was there any attempt to allow for load carried, as has been insisted upon by Toops (16), since the enrollments in the college were quite uniform in this respect.

Table 2 shows the coefficients of correlation of zero order. Attention is called here to the correlation represented between

TABLE 2

Correlation coefficients of zero order tentatively used in predicting advanced scholastic success of 107 engineering students

	1	2	3	4	5	6	7	8
1. Adv. engin. (criterion)		.584	.276	.830	.453	.356	.611	.400
2. Fresh. engin.	.584		.383	.535	.400	.466	.738	.282
3. Fresh. English	.276	.383		.518	.378	.459	.781	.288
4. Fresh. math.	.630	.535	.518		.379	.393	.816	.420
5. Fresh. mechan. draw.	.453	.400	.378	.379		.289	.573	.406
6. Fresh. chem.	.356	.466	.459	.393	.289		.680	.084
7. Gen. fresh. ave.	.611	.738	.781	.816	.573	.680		.382
8. Otis score	.400	.282	.288	.420	.406	.084	.382	

the various independent variables and the criterion. When these variables are set down in order of the values of their coefficients they are as follows: mathematics ($r_{14} = .630$), freshman general average ($r_{17} = .611$), introductory engineering ($r_{12} = .584$), mechanical drawing ($r_{15} = .453$), Otis ($r_{18} = .400$), chemistry ($r_{16} = .356$), and English ($r_{13} = .276$).

These coefficients, of course, are too low for effective prognosis; however, they compare favorably with those of most measures of scholarship. Mathematics stands out as a single prognostic measure; it is slightly superior to the combined freshman courses. Next comes the introductory engineering course. If it were offered during the first semester instead of the second as it now is it would give a fair prognosis of the student's ability at the end of the first half year's work. Then follows mechanical drawing. It is rather surprising that this coefficient is as high as it is. Next is the Otis score which is considerably lower than was hoped for. However, this measure of ability which requires about seventy-five minutes to administer is superior to chemistry and English, the next in order, which cover a whole year's work.

The reliability of these measures excepting Otis is very low, ranging from .413 to .588. The coefficient of reliability for the Otis test, which is .967, was furnished by the author of the test, that of all the others, excepting introductory engineering, by correlating first semester grades with second semester grades in the same department. Since the engineering course is offered only one semester no technique for determining its reliability was readily available.

In order to reduce the endless labor of computing the partials for all the independent variables, chemistry and English, whose combined effect upon R would be likely very small anyway, are arbitrarily omitted. So the basis of prediction is a battery made up of the following: engineering, mathematics, mechanical drawing, and Otis score. The coefficients of correlation of the third order are presented in Table 3.

TABLE 3

Coefficients of correlation of the third order used in predicting advanced scholastic success of 107 engineering students. (Variables: (1) advanced engineering, (2) freshman engineering, (4) freshman mathematics, (5) freshman mechanical drawing, (8) Otis score)

$r_{12.458} = -.032$	$r_{19.345} = .172$	$r_{48.125} = .212$
$r_{14.258} = .484$	$r_{28.145} = .031$	$r_{50.124} = .254$
$r_{15.248} = .005$		

Some coefficients which are not needed in deriving the regression equation are included. The corresponding standard deviations are shown in Table 4, and the regression coefficients or weights in Table 5.

TABLE 4

Standard deviations of the third order used in predicting advanced scholastic success of 107 engineering students. (Variables: (1) advanced engineering, (2) freshman engineering, (4) freshman mathematics, (5) freshman mechanical drawing, (8) Otis score)

$\sigma 1.2458 = .675$	$\sigma 5.1248 = .997$
$\sigma 2.1458 = 1.097$	$\sigma 8.1245 = 20.622$
$\sigma 4.1258 = .993$	

TABLE 5

Weights of variables used in predicting advanced scholastic success of 107 engineering students. (Variables: (1) advanced engineering, (2) freshman engineering, (4) freshman mathematics, (5) freshman mechanical drawing, (8) Otis score)

$b_{12.458} = -.019$	$b_{15.248} = .0035$
$b_{14.258} = .329$	$b_{18.245} = .056$

The multiple coefficient of the combination chosen is $R_{1(2458)} = .690$, the standard error of estimate is .675, and the efficiency of prediction is 28. It will be recalled that the coefficient of correlation between mathematics and the criterion was .630 (Table 2), which has a predictive efficiency

of 22 per cent. All of the labor spent in determining the above R results in raising the efficiency of the predicting device from 22 percent to only 28 percent above that of chance prediction. The R might have been slightly higher if the coefficients of correlation had been corrected for attenuation. This was not done, however, since the method of computing some of the coefficients of reliability was at best rather crude and probably, therefore, would not have justified the additional labor of refinement. However, the raw R is on the upper border of that range of coefficients (.60 to .70) which, according to Hull (4), indicate "genuine but limited value" in prognosis.

RESULTS

The regression equation. Since the variables selected for prognosis are (2) engineering, (4) mathematics, (5) mechanical drawing, and (8) Otis score, the regression equation is:

$$X_1 = b_{12\cdot 458}X_2 + b_{14\cdot 258}X_4 + b_{15\cdot 248}X_5 + b_{18\cdot 245}X_8 - K \quad \text{..... (1)}$$

in which $\overline{X}_1$ is the honor points in advanced engineering of a given student to be predicted, X_2, X_4, X_5, and X_8 are his observed honor points in introductory engineering, mathematics, mechanical drawing, and his score on the Otis test respectively, and the b's are the weights to be given to each. When the weights and the constant term for the engineering data are substituted, equation (1) becomes:

$$\overline{X} = -.019X_2 + .329X_4 + .0035X_5 + .056X_8 - 5.995 \quad \text{..... (2).}$$

In order to find the predicted honor points in advanced engineering of a given student at the end of his freshman year, it is necessary only to substitute the appropriate freshman honor points and the Otis score and then perform the operations indicated.

The predicted scores. In Fig. 1 are represented the predicted honor points (y-axis) plotted against the observed honor points (x-axis) of the advanced engineering courses for the total group. At the outset attention is called to the discrepancy between the ranges of honor points in the two arrays.

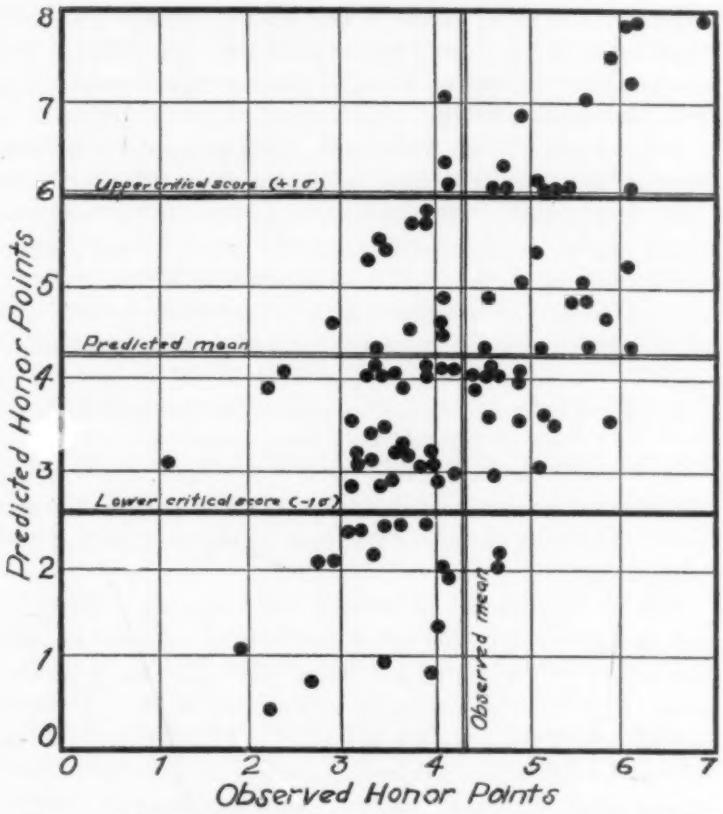


FIG. 1. Relation of predicted honor points to observed honor points in advanced engineering courses of the University of Oklahoma. Independent variables are honor points in freshman engineering, freshman mathematics, freshman mechanical drawing, and Otis intelligence scores. N = 107, R = .69, E = 28).

It will be recalled that the highest score obtainable in a course was to be seven honor points, but in Fig. 1 the predicted scores range to eight. This apparent discrepancy can be explained by the fact that any predicted score is a given amount plus or minus a certain error of estimate and by the fact that the

upper limit of seven points in the grading system is an arbitrary limit which cannot actually be exceeded although certain students theoretically might deserve higher markings if the scale permitted them.

Critical scores. An attempt is made to apply a modified form of Thurstone's critical score technique (15) to the results of this study. One difference is that in his study observed scores (instructors' estimates) were compared with scores made on the prognostic tests, whereas in this study the observed scores are compared with the *predicted* scores. Another difference is in the basis of determining the critical scores.

In Fig. 1 lines are drawn at certain crucial points on the ordinate, namely, at $\frac{X-M}{\sigma}=+1$, and at $\frac{X-M}{\sigma}=-1$. Since the standard deviation of the distribution of predicted grades is $\pm$ 1.65 and the mean is 4.24, these points are 5.89 and 2.59, respectively.

Now let us see what is revealed when the predicted scores and the observed scores are considered in relation to these crucial points. First of all, let us note that the predicted mean (4.24 honor points) and the observed mean (4.36 honor points) are almost exactly the same. Furthermore, of the eighteen students whose predicted scores are below the lower critical point, all except two are below the mean of observed honor points. Of the eighteen whose predicted honor points are above the upper critical point, all except three are above the observed mean. In other words, of the thirty-six exceptional students in the group, thirty-one (86 per cent) could have been somewhat definitely placed at the end of the freshman year with respect to their ability to pursue advanced courses in engineering. Of the seven students who withdrew from the College, apparently on account of poor scholarship, all except one are below the lower critical point and all are below the predicted mean of honor points.

USE OF THE PREDICTIVE DEVICE

The use which may be made of this prognostic device depends upon the purpose of the one employing it. If it is to be used to limit the enrollment in advanced classes to those who will be almost certain to achieve scholarship superior to a general average, then the predicted scores of such prospective students at the end of the freshman year should equal or exceed the lower critical score. The student who makes a lower predicted mark is quite certain to make inferior grades if actually admitted to such advanced courses.

If, in a similar way, the purpose is to locate those students at the end of the freshman year who are expected to have difficulty in completing the advanced work, then the students sought will again be those whose predicted grades lie somewhere below the lower critical point. In these cases the device can become diagnostic as well as prognostic. The particular freshman subjects in which low grades have been made may be assigned for review or reenrollment, thereby strengthening the basic work. This should be particularly effective in cases where the Otis scores are average or above.

If it is desired, on the other hand, to predetermine the potentially superior students, say, for election to scholarships and undergraduate assistantships, that group whose predicted marks equal or exceed the upper critical point are, with a fair degree of certainty, the students sought.

SUMMARY AND CONCLUSIONS

This is the report of a study in predicting success in the College of Engineering of the University of Oklahoma, based on the records of one hundred seven students. The criterion is scholarship, as represented by honor points, in courses in advanced engineering on the undergraduate level. The independent variables are the honor points gained in certain freshman engineering courses, namely, mathematics, introductory engineering and mechanical drawing, and the score on the Otis Advanced Intelligence Scale. The coefficients of correlation

between these scores and those of the criterion are mathematics, $r_{14} = .630$; introductory engineering, $r_{12} = .584$; mechanical drawing, $r_{15} = .453$; and Otis, $r_{18} = .400$. The coefficients of reliability of the freshman subjects used range from .413 to .588 and that for the Otis test is .967. The multiple R is $R_{1(2458)} = .690$ and its efficiency is 28.

A kind of critical score technique is used in comparing the predicted scores with the observed scores. Of thirty-six exceptional students (that is, students whose honor points lie above or below the critical points) 86 per cent could have been fairly definitely placed with respect to their potential ability in advanced engineering courses at the end of the freshman year. Of the seven who withdrew from the college, apparently from poor scholarship, all are below the predicted average and all except one are below the lower critical point in predicted grades.

It may be said in conclusion that the score on the Otis Advanced Intelligence Scale and honor points in freshman prescribed courses constitute only a fairly accurate basis for determining potential ability of students in advanced engineering courses.

REFERENCES

1. Asher, E. J. The predictive value of mental tests that satisfy Spearman's tetrad criterion. J. Appl. Psychol., 1929, 13, 152–158.
2. Groves, T. Freshman grades as a predictive basis for success in law courses. Unpublished M.A. Thesis, University of Oklahoma Library. 1926.
3. Holzinger, K. J. Statistical résumé of the Spearman two-factor theory. University of Chicago Press, 1930, 16–21.
4. Hull, C. L. The correlation coefficient and its prognostic significance. J. Educ. Res., 1927, 15, 327–338.
5. Hull, C. L., and Limp, C. E. The differentiation of the aptitudes of an individual by means of test batteries. J. Educ. Psychol. 1925, 16, 73–88.
6. Lewis, F. C. A study in formal discipline. School Rev., 1905, 13, 281–292.

7. LIMP, C. E. The use of the regression equation in determining the aptitudes of an individual. J. Educ. Psychol., 1925, 16, 414–418.
8. MAY, M. A. Predicting academic success. J. Educ. Psychol., 1923, 14, 429–440.
9. MCDONOUGH, R. General factor in a table of intercorrelations. Science, 1929, 69, 402.
10. NELSON, M. J., and DENNY, E. C. The Terman and Thurstone group tests as criteria for predicting college success. School and Soc., 1927, 26, 501–502.
11. PIERSON, C. D., and NETTELS, C. H. A study of high school seniors to determine who shall be recommended to college. School and Soc., 1928, 28, 215–216.
12. RECTOR, W. G. A study in the prediction of high school success. J. Educ. Psychol., 1925, 16, 28–37.
13. RIETZ, H. L. On the correlation of the marks of students in mathematics and in law. J. Educ. Psychol., 1916, 7, 87–92.
14. SPEARMAN, C. The abilities of man. Macmillan, 1927.
15. THURSTONE, L. L. Mental tests for college entrance. J. Educ. Psychol., 1919, 10, 129–142.
16. TOOPS, H. A. The prediction of scholastic success in college. School and Soc., 1927, 25, 265–268.

SUGGESTIBILITY AND THE FORM OF THE QUESTION

HAROLD E. BURTT
Ohio State University
AND
HAROLD V. GASKILL
Iowa State College

PROBLEM

It is well known that suggestibility is affected considerably by the form in which a question is asked. The earlier "Aussage" experiments brought out the greater suggestiveness of the expectative type of question, such as "Was there dog in the picture?" and the still greater suggestiveness of the implicative form such as, "Was the dog black or white?" when no dog was present at all. Quantitative measures of the degree of suggestiveness produced by such questions have been reported. A more comprehensive study was made by Muscio.* He gave his subjects short bits of moving picture film of about 30 seconds' duration, and then asked questions about them systematically, using eight different forms, as follows:

A. Did you see a .....
B. Did you see the .....
C. Didn't you see a .....
D. Didn't you see the .....
W. Was there a .....
X. Wasn't there a .....
Y. Was the M K?
Z. Was the K, M or N?

After showing each film for two or more times, until the subject was fairly familiar with it, the questions were asked in the above forms and results tabulated to indicate the extent

* Muscio, B. The Influence of the Form of Question. *British Journal of Psychology*, 1915, 8, 351–389.

of suggestion. Muscio, however, construed suggestion to mean any admission of the presence of an object. This construction included correct observation and report as well as incorrect replies due to the suggestiveness of the question. The present writers feel that this is not the interpretation that is of most importance in actual practice. From the standpoint of testimony, for example, we are particularly concerned with the cases in which the subject will erroneously admit the presence of an object due to the way in which the question is asked. We are inclined to construe suggestion from this standpoint to mean admissions which are actually incorrect. Muscio came to the conclusion that there was greater suggestiveness with the indefinite article "a," with the negative and with the subjective form in which the personal pronoun "you" occurred. We shall see later that perhaps Muscio's results need to be somewhat reinterpreted in the light of the foregoing issue as to whether admission of the presence of objects which actually were present should be construed as suggestion or not. We will recur to this point.

MATERIAL

The materials used in the present experiment consisted of two moving picture news reels of about twelve minutes' duration each. They were reels which had not been shown for several months in local theaters so that there was no effect of very recent experience involved. Both were silent and had the usual printed headings at frequent intervals. The first of these which we will call "H" involved the following scenes: American sailors visiting an oriental village; harvesting ice; wrecked airplane; cross-country race; automobile race; motorcycle police doing stunts; army kitchen; Swiss army manoeuvering on skiis; toboggan slide.

The second reel (K) comprised the following scenes: dog team races; overhauling navy airplane carrier; aviator who was lost subsequently in Siberia; airplane that flapped wings; riot in New York; Japanese field day and fire apparatus dem-

onstration; fishing vessel with frozen rigging; marble quarry; skating races.

METHOD

The general procedure was to show the film once, and immediately question the subjects about details in the film. The reel was projected in a darkened room and the lights were turned on during the subsequent questions. Mimeographed blanks numbered with digits from one to thirty-six were provided and the subjects merely wrote "Yes" or "No" after each number as the questions were given verbally. The two experimenters alternated in giving the questions. Usually each group of subjects was shown both films, H and K, at the same sitting. Film H always preceded the other.

The subjects were simply instructed that this was an experiment on observation, that they were to be shown a moving picture news reel and would subsequently be questioned about it. In most of the series they were instructed to answer either "Yes" or "No" for each question. A few series were introduced in which attention was called to the possibility of writing a question mark if they were uncertain. Of course some question marks occurred in the series where the instructions were to answer merely "Yes" or "No," but these were very few indeed. The results for the two types of instruction were of course analyzed separately.

Interest in the present experiment was confined to six wordings of the question—the first six used by Muscio. The same notation will be followed, namely:

A. Did you see a .....
B. Did you see the .....
C. Didn't you see a .....
D. Didn't you see the .....
W. Was there a .....
X. Wasn't there a .....

For each film 30 items of information were selected and questions were asked about them, in these six different forms.

The questions were all worded in such a way that a "Yes" answer would be wrong and would indicate to a certain extent that the suggestion had been effective. In addition six other questions were used as fillers, in which a "Yes" answer was correct. These six questions were made fairly obvious so that the subject would realize that there were some "Yes" answers and would not develop a "set" for "No" answers. These six filling questions were the same in every series for a given film. The other thirty questions, however, were stacked systematically. Six series of questions about a film were arranged so that each item was interrogated once in each of the six wordings of the question. The items were always kept in the same order; that is, the second question was always about the sailor going down the gang-plank, but sometimes the question read, "Did you see a sailor on the gang-plank?" at other times, "Didn't you see the sailor on the gang-plank?" or "Was there a sailor on the gang-plank?" etc. This order of questions was arranged from a master table, such that if item 2 was in one series interrogated in form A in the next series it would be in form B, in the next form C, etc. Averaging the results of the six series for a given wording of the question should thus iron out errors due to the difficulty of a particular item. It naturally was impossible to get items of equal difficulty and seemed best to make allowance in this way. However the results were also analyzed directly with reference to the empirical difficulty of the items.

All six series of questions were used for each film. It was impossible to keep the numbers of subjects absolutely the same, because they had to be taken in the groups available in the elementary psychology sections. The subjects were, with the exception of a group of fourth grade students, members of classes in elementary psychology at the Ohio State University in the spring of 1930. This project replaced one of their usual classroom exercises.

Altogether there were twenty-five different series performed, with groups of subjects varying from twenty-nine to

fifty-eight in number. There were thirteen series with film H and twelve with film K. Nineteen of the series involved instructions to answer "Yes" and "No." Eight hundred eighteen subjects participated in these series. In six other series involving two hundred seventy-two subjects the instructions were to answer "Yes," "No" or "Don't know." This supplementary series was introduced particularly in order to analyze "caution" and the results are discussed separately. The proportion of "Yes" and "No" answers in the supplementary series may be compared with that in the main series, and then the proportion of "Don't know" answers may be analyzed separately.

RESULTS

Every question in a given series was scored according to whether the answer was "Yes" or "No" and then according to the number of subjects who indicated "Yes," "No," or "Don't know." These figures were also expressed as the percent of the subjects answering "Yes," "No" or "Don't know." The original data simply gave these numbers and percents for every item or question. It will be recalled, however, that there were always five questions of a certain form such as A, B, C, (supra) in a given series and there were six series of questions for each film. The next step then was to assemble all the data for questions in form A, form B, etc. This was facilitated by a stencil which showed for each series the location of each form of question.

The results are analyzed from several angles in Table I. The various methods of analysis will be explained and then they will be discussed together. The figures in the left block are for the main experiments in which the "yes-no" instructions were given. The middle block analyzes the "Yes" answers in the supplementary experiment where the "don't know" alternative was included in the instructions. The actual analysis of the "don't know" answers, *i.e.*, "caution," is made in the right block.

TABLE I

	MAIN SERIES						SUPPLEMENTARY SERIES			
							"Yes" Answers		"Don't know" Answers	
Form of question	Gross per cent of total	Av. per cent	Deviate from difficulty gross	Weighted deviation	Median of deviates	Per cent plus deviations	Gross per cent of total	Weighted deviation	Gross per cent of total	Weighted deviation
A	17.5	17.0	−1.40	− .94	−1.86	34	19.5	.20	4.1	−15.4
B	18.2	17.8	− .70	− .50	−1.75	37	14.6	−1.70	4.3	−11.8
C	17.6	18.4	− .55	− .61	−1.74	33	14.3	−1.53	4.2	−11.6
D	16.8	16.4	−1.49	−1.09	−1.42	33	18.8	− .88	2.4	−17.3
W	18.4	17.8	.70	.89	− .75	37	22.8	2.18	16.0	− 4.7
X	19.1	19.7	1.45	1.71	− .63	44	17.3	1.13	19.3	3.1

A computation was first made of the grand total number of "Yes" answers to all questions on both films that were asked in a particular form such as A. These numbers were reduced to a per cent of the total number of answers,—4,090 in the main experiment. The successive rows of the table give results for the six forms of question as indicated by the letters in the first column. The second column gives the above-mentioned per cents. For example, out of the 4,090 replies to a question in form A ("Did you see a.....") 17.5 per cent of the replies were "Yes," *i.e.*, incorrect. Similarly of the 4,090 answers to questions in form B ("Did you see the.....") 18.2 per cent were "Yes."

Another procedure involved computing for each series and each group of subjects the proportion of subjects who answered "Yes" on a given item. These per cents were then assembled for all form A questions, form B questions, etc., and then the average of all these individual per cents taken for the entire group of form A questions. These figures are given in column 3. For instance the average of all the per-cents of yes answers to questions in form A is 17.0; the average of the per cents for form B is 17.8, etc.

Each form of question was given an equal number of times and each item was interrogated equally often in each form. It was not possible always to have the same number of subjects answer each item in each form and consequently there might be a possible source of error, due to the difficulty of the item. Some of the items naturally were more difficult than others,—for example, the presence of a policeman at a riot would probably be more correctly reported than the name of the ten-cent store in the background. If by any chance the difficult items were interrogated to a larger number of subjects in a certain form we might get an erroneous indication of suggestiveness of that form. While a large error of this sort is not very probable, it seemed wise to analyze the results so as to obviate it. This was done by computing the difficulty of each particular item. For instance, we tabulated the number of subjects who said "Yes" to a particular item, such as the presence of the sailor on the gang-plank, regardless of the series or regardless of the form of the question. We simply obtained a grand total of the number of subjects who said "Yes" with reference to a particular item and then reduced this to a percent of the total number of subjects. This gave an indication of the difficulty of that item, because the larger the percent of "Yes" answers, that is the larger percent of incorrect replies, the more difficult might that item be considered to be. For instance with item 2 in film H, 10 per cent of all the subjects marked "Yes." This percent indicates the difficulty of that item. With such indices of difficulty available, the original data were gone over and in every question in every series the percent of "Yes" answers (the type of data used in obtaining column 3) was reduced to deviation from the difficulty of the item. If, for example, the general difficulty of the item about the gang-plank was 17 per cent and in a particular series and form of question, 20 per cent of the subjects said "Yes" this represented a deviation of +3, whereas, if 16 per cent of them said "Yes," this represented a deviation of −1. Obviously the larger the fig-

ure in a plus direction the greater the suggestion. These deviations were then totalled algebraically and an average obtained for all questions of a given form. The averages occur in column 4. For instance, on the average in Form A the percent of "Yes" answers was about 1.4 per cent less than the per cent of "Yes" answers made on the same item regardless of the form of question.

Inasmuch as some series involved more subjects than others, it seemed advisable further to take the data that were used in obtaining column 4 and weight the deviations according to the number of subjects. Each deviation was multiplied by the number of subjects involved in producing it and these totals were divided by the grand total 4,090 to get the weighted deviation. These weighted deviations appear in column 5.

While the distributions of the deviations contributing to the averages in column 4 seemed fairly normal, it was thought best to compute the median of these deviations also. These medians appear in column 6.

It was also interesting to note for all the data for a given form of question the proportion of deviations that were greater than the difficulty of the item and *vice versa*. In column 7 we have the per cent of the deviations that were plus, *i.e.*, that involved a larger percent of "Yes" answers than the item in question showed on the average. In this connection there were quite a few deviations of zero, so that these results might have to be somewhat qualified. In form D, for example, there were 10 per cent of these zero values.

In the next two columns we have similar results for the supplementary series in which the instructions were to answer "Yes," "No" or "Don't know." The results in this block pertain only to the "Yes" answers. It did not seem profitable to compute the results by the six methods just described but only by those methods that seemed most valuable in the first instance. With reference to this latter point we note that the data in columns 2, 4 and 5 seem to show a fair

amount of agreement. If we simply take the rank order of the six forms of question there is a perfect correlation between columns 2 and 4, and high ones between 2 and 5 and between 4 and 5. The remaining columns do not seem quite so typical. Consequently in column 8 the results are presented from the standpoint of the gross per cent of the total replies which were "Yes" for a given form of question. These figures were computed in exactly the same way as the figures in column 2. The ninth column is identical in form with column 5.

The "Don't know" answers in this supplementary series are analyzed in columns 10 and 11. The method of analysis is the same as that used in columns 2 and 5 respectively. In column 10 we have the gross per cent of "Don't know" answers and in column 11 these are transmuted into terms of deviations from the difficulty of each item and then these deviations weighted. This latter procedure might perhaps be somewhat less significant because the difficulty of the items as indicated by the average per cent errors often was rather widely different from the proportion of "Don't know" answers. However, the data were computed in this way, just as in column 9.

It is obvious that there are some differences between the various forms of question. Many of these differences are slight but nevertheless the number of cases is rather large. The only way to interpret the results adequately is to compute the standard deviation of the differences. For columns 2, 8 and 10 this involved the formula for the standard deviation of a per cent. For column 6 it was computed from the quartile deviation and in the remaining columns from the standard deviation of the original distribution. For column 7 it did not seem profitable to compute it at all.

In Table II the results are presented from the standpoint of three different variations in the form of the question, and the differences between the various forms are interpreted relative to the standard deviation of the differences. The successive blocks of the table going from left to right are the same as in

Table I. The upper portion of the table analyzes the results with reference to the effect of the article, the middle portion the effect of the negative *versus* positive form, and the lowest portion the subjective *versus* objective form of question.

In the first row of the table which compares forms A and B, we note that from the standpoint of the gross per cent of totals B shows more suggestiveness than A. In every instance the letter at the left of the diagonal line indicates greater suggestiveness for that type of question. Directly below these letters we note that the actual per cent is 18.2 for B and 17.5 for A. These per cents are taken directly from Table I. At the right of this we note that the difference between these per cents divided by the standard deviation of the difference is .83. Reading across the row we have all the other comparisons of Forms A and B for the actual per cents or other mea-

TABLE II
MAIN SERIES

Form of question	Gross per cent of total	Average of per cent	Deviate from difficulty gross	Weighted deviation	Median of deviates	Per cent plus deviations
A–B	B/A .83 18.2/17.5	B/A .25 17.8/17.0	B/A .66 –.70/–1.40	B/A .46 –.50/ –.94	B/A .10 –1.75/–1.86	B/A 37/34
C–D	C/D 1.00 17.6/16.8	C/D .66 18.4/16.4	C/D .74 –.55/–1.49	C/D .41 –.61/–1.09	D/C .27 –1.42/–1.74	C/D 33/33
A–C	C/A .12 17.6/17.5	C/A .43 18.4/17.0	C/A .72 –.55/–1.40	C/A .34 –.61/ –.94	C/A .11 –1.74/–1.86	A/C 34/33
B–D	B/D 1.67 18.2/16.8	B/D .46 17.8/16.4	B/D .66 –.70/–1.49	B/D .51 –.50/–1.09	D/B .27 –1.42/–1.75	B/D 37/33
W–X	X/W .81 19.1/18.4	X/W .58 19.7/17.8	X/W .63 1.45/ .70	X/W .73 1.71/ .89	X/W .10 –.63/ –.75	X/W 44/37
A–W	W/A 1.06 18.4/17.5	W/A .25 17.8/17.0	W/A 1.87 .70/–1.40	W/A 1.90 .89/ –.94	W/A .98 –.75/–1.86	W/A 37/34
C–X	X/C 1.75 19.1/17.6	X/C .39 19.7/18.4	X/C 1.56 1.45/ –.55	X/C 2.09 1.71/ –.61	X/C .92 –.63/–1.74	X/C 44/33

TABLE II (*Continued*)
SUPPLEMENTARY SERIES

Form of question	"YES" ANSWERS		"DON'T KNOW" ANSWERS	
	Gross per cent of total	Weighted deviation	Gross per cent of total	Weighted deviation
A–B	A/B 3.50 19.5/14.6	A/B .91 0.20/–1.70	B/A .26 4.3/4.1	B/A .82 –11.8/–15.4
C–D	D/C 3.10 18.8/14.3	D/C .29 –.88/–1.53	C/D 2.40 4.2/2.4	C/D 1.10 –11.6/–17.3
A–C	A/C 3.60 19.5/14.3	A/C .84 .20/–1.53	C/A .13 4.2/4.1	C/A .87 –11.6/–15.4
B–D	D/B 2.90 18.8/14.6	D/B .36 –.88/–1.70	B/D 2.70 4.3/2.4	B/D 1.06 –11.8/–17.3
W–X	W/X 3.60 22.8/17.3	W/X .62 2.18/ 1.13	X/W 2.26 19.3/16.0	X/W 1.21 3.1/–4.7
A–W	W/A 2.10 22.8/19.5	W/A .93 2.18/ .20	W/A 10.50 16.0/4.1	W/A 1.65 –4.7/–15.4
C–X	X/C 2.10 17.3/14.3	X/C 1.64 1.13/–1.53	X/C 12.50 19.3/4.2	X/C 3.35 3.1/–11.6

sures of suggestibility and the difference divided by the standard deviation of the difference. The next section below gives similar results for Forms C and D. Below this we have other comparisons as indicated.

The most significant things in the table, of course, are these ratios of the differences to their standard deviations. Certainly if this ratio is less than one, it should receive very little consideration because under those circumstances there are only 84 chances out of 100 that the difference is a real one. If this difference is 1.5 the chances are 93 out of 100, and if it is 2, 98 out of 100 that the difference is significant. We may thus analyze the suggestibility as brought out in both the main series and the supplementary series.

Turning first to the form of the article it will be recalled that forms A and C embodied the indefinite article "a," whereas B and D involved the definite article "the." In the main series B shows more suggestiveness than A in every method of analysis but none of the differences are statistically significant. In the supplementary series the gross per cents indicate greater suggestiveness for A rather than B, and this difference is significant. In the other method of analysis the difference is not significant. With reference to forms C and D, which both involve the negative, we note by most methods of analysis that in the main series C shows more suggestiveness than B, but only one of these differences is of any possible significance. In the supplementary series one set of figures tends to reverse the general trend of the main series by a significant difference, while the other is insignificant. Again with reference to "caution" there is one statistically significant result of more "Don't know" answers for C than for D but the other comparisons are equivocal. On the whole it would seem that the results with reference to the form of the article as effecting suggestibility are rather equivocal. It is doubtful if there is a sufficiently uniform trend to be of any practical value.

Turning now to the analysis of the negatives, with questions A and C, which both involve the indefinite article, but one in the negative and the other the positive form, most of the results in the main series indicate greater suggestiveness for C. The differences are all insignificant, however. In the supplementary series one set of figures reverses this, giving A a significant superiority to C, but the other is insignificant. With forms B and D, which differ in the negative, but both involve the definite article, the weight of evidence in the main series indicates that B carries more suggestiveness. One of these differences is probably significant. On the other hand, in the supplementary series the only significant difference reverses this trend in the main series. Finally with questions X and W, which are both in the objective form and involve

the indefinite article but vary in the use of the negatives, X seems more suggestive than W throughout the main series, but the differences are all less than the standard deviation. In the supplementary series, however, the per cent of the total "Yes" answers is significantly greater for W than for X. These results are somewhat difficult to interpret. As far as "Yes" answers go, it would appear that the supplementary series contradicts the main series on any cases where the differences are statistically significant. We may note that in the main series all the different methods of analysis yield substantially the same results while the two methods of analysis in the supplementary series also agree with one another. We may note further that the "Don't know" answers in the supplementary series seem to bear an inverse relation to the "Yes" answers in that series. For instance, while form A has more "Yes" answers it has less caution, and while D has more "Yes" replies, its caution is less than that of B. Similarly, W gets more "Yes" replies and fewer "Don't know" replies. The trend for the "Don't know" answers corresponds exactly to the trend for "Yes" answers in the main series. As far as forms A, C, X and W are concerned we might be tempted to suggest that when the subject is compelled to give a categorical "Yes" or "No" answer the negative form of question leads him to say "Yes" erroneously, but when he is allowed to give a "Yes" or "No" answer, or to say that he does not know, then the negative form tends to produce more caution, so that he will more frequently say that he doesn't know and consequently make a smaller proportion of incorrect "Yes" replies. However, forms B and D do not agree with this finding. When the questions are in the subjective form with the definite article, it appears that the positive form has more suggestibility when categorical answers are demanded and more caution when the "Don't know" possibility is open. It would perhaps be advisable to suspend judgment on this point pending further experiments. It is interesting to note, however, that in Muscio's results, if

we take the per cent of cases in which a "Yes" answer was given, when the content was false,—that is, construing the suggestiveness in our present sense,—his data give relations like those in column 2 of our table in every case.

With reference to the subjective-objective form of the question the results are fairly clear. In the main series the majority of the differences are significant and the different methods of analysis indicate without exception a greater suggestiveness of Form W *versus* A, and of X *versus* C. Both of these forms involve the objective statement. Similarly in the supplementary series both methods of analysis indicate a greater suggestiveness for W and X,—three of the four differences being of some statistical significance. It will be recalled that this finding is opposed to that of Muscio. However, as above suggested he interprets suggestiveness to mean the admission of the presence of an object whether or not the admission was correct. We have construed it only in the sense of admitting erroneous statements, which construction we feel is closer to the actual practical implications of the problem. Incidentally if we consider in Muscio's original data only the cases in which false content was involved, while A shows a little more suggestiveness than W, X on the other hand shows greater suggestiveness than D. His results from that standpoint then might be somewhat equivocal whereas ours show very consistently the greater suggestiveness of the objective form of question. Perhaps when we question a person and make it somewhat of a personal matter as to whether or not he saw some particular object, he is a little less inclined to erroneously say "Yes" than if the question is asked somewhat more impersonally. We note also, however, that the objective form for some reason manifests greater caution and these differences are presumably significant. We do not find a trend as in the case of the negative just cited where a certain form of question produces greater caution and hence less suggestiveness. In the present instance the objective form not only leads the person to make more mistakes but also to

give more equivocal answers. Obviously in getting a correct statement of fact the advantage lies with the subjective form of the question. We are less apt to get an erroneous statement and also less apt to have the subject hesitate to answer at all. This seems to be the clearest result in the study.

INCIDENTAL PROBLEMS

The preceding results regarding caution and tendency to give "Yes" answers suggests the possibility of investigating the relation between these two. Consequently a correlation was run between per cent "Yes" answers and per cent "Don't know" answers for all the individuals in the supplementary series. This correlation is .06. It might have been anticipated in view of the fact that there was apparently some inverse relation between caution and suggestiveness with reference to the negative forms of question and a direct relation with reference to the positive forms of question. It would seem that if there is any relation between the two it varies with the form of the question rather than representing any general tendency for caution and suggestiveness to have a direct or inverse relation.

Two experimenters participated in giving the questions,—on different series of course. It seemed worth while to determine whether one examiner had any greater tendency to get "Yes" replies than did the other. This might have a bearing on the question of this tendency of different prosecutors or different lawyers to get more or less "Yes" replies to suggestive questions. The senior examiner was some fifteen years older than the junior and to get a rough notion of the trend the per cent "Yes" replies in the main series were tabulated separately for the two examiners. There was a slight difference in favor of the senior member. With his subjects 19.3 per cent of the suggestive questions received a "Yes" answer, and with the other 16.1 per cent. The difference between these per cents is 6.7 times the standard deviation of the difference.

A group of twenty subjects were children from the fourth grade. It is interesting to note that in general their per cent of "Yes" answers was 55, whereas the per cent for the other subjects who took the same series of questions was about 18. This corresponds to the usual finding of much greater suggestibility for children.

SUMMARY

A moving picture news reel was shown and the subjects then questioned verbally about details in the picture. Six forms of question were used in order to evaluate the definite *versus* indefinite article, the negative *versus* the positive form and the subjective *versus* the objective form. The questions were answered on a mimeographed blank by writing "Yes," "No" or "Don't know." The proportion of erroneous "Yes" answers was assumed to be an indication of suggestiveness as affected by the form of question. Upwards of 5000 answers to each form of question were available for analysis.

The results for definite *versus* indefinite article are equivocal and no conclusions are warranted. There seems somewhat of a trend for the negative to cause greater suggestiveness when categorical answers are demanded and otherwise greater caution but this trend is contradicted in the comparison of questions in the subjective form with the definite article.

The objective form of question shows clearly a greater suggestiveness and also a higher degree of caution.

PROFESSOR H. F. ADAMS AND THE TWO-FACTOR THEORY OF ABILITY

ROBERT H. THOULESS
University of Glasgow

In the course of an investigation into the difference between objective and subjective tests,[1] Professor H. F. Adams was led to believe that his work rendered untenable Professor Spearman's two-factor theory of ability.[2,3] It would seem indeed to have an even more revolutionary effect, for, if Adams were right, it would be necessary to give up altogether the idea that a correlation between test results was any measure of a relationship between abilities.

It is with this part only of Adams's work that this paper is concerned and not at all with his method of distinguishing between subjective and objective tests. The argument with respect to the theory of general ability is based on objective test results and it is to these only that this paper will refer.

Professor Adams's argument may be summarized as follows. Spearman's evidence for the existence of general ability rests on the truth of a series of equations connecting correlations between test results. The same equations are found by Adams to connect correlations between test results, in which it is the standard or correct order of test material which is found to play the part attributed by Spearman to *g*.

[1] "An Objectivity-subjectivity Ratio for Scales of Measurement," H. F. Adams, *The Journal of Social Psychology*, I, 1930, pp. 122–135.

[2] "The Theory of Two Factors: an alternative explanation. Part I. General Factor, 'G'," H. F. Adams, THE JOURNAL OF APPLIED PSYCHOLOGY, XV, 1931, pp. 16–34.

[3] "A Theory of Two Factors: an alternative explanation. Part II. Specific Factors," H. F. Adams, THE JOURNAL OF APPLIED PSYCHOLOGY, XV, 1931, pp. 358–377.

Therefore g is identical with the standard and is not a general ability.

Adams finds, for example, that the correlation between two tests, a and b, is exactly the product of the correlation of each with its standard. In this case, the correlation between the two tests can hardly be (as is commonly supposed) a measure of the relationship between the two abilities tested.

The reader may be perplexed to understand what is meant by the correlation of a test with its standard. This is clearly explained by the author and we discover from this explanation that although Professor Adams uses such familiar terms as "correlation between two tests" and "self-correlation of a test," and the familiar symbols r_{ab}, etc., he is using these terms and symbols in a sense entirely different from that of Spearman and of most other psychologists, and that Adams's r_{ab} stands for a relationship which all will agree is not a relationship between abilities.

It is very necessary to understand exactly what is the meaning of r_{ab} in Adams's paper. A subject is given a series of test articles, such as ten weights forming a geometrical series of increasing weight, and is asked to arrange them in order of weight. Unless he is able perfectly to discriminate the test weights, the order he arranges them in will not coincide with the standard order. A rank-difference correlation is worked out between the order he produces and the standard order. The r_{ax} of Adams is the mean of the correlations thus found for a group of subjects: his r_{ab} is the mean of the correlations similarly obtained between the orderings for each individual subject in two different tests.

This is clearly a totally different meaning for the symbol r_{ab} from that used in the work of Spearman and other investigators on correlations between abilities, in which a group of testees is scored or arranged in two different tests, and r_{ab} is the correlation between these two sets of scores. The vital difference between this and Adams's symbol is that this represents a relationship between abilities whereas Adams's does

not. This r_{ab} will be high if the two abilities are similar or identical and low if they are very dissimilar. Adams's r_{ab} is not a relationship between abilities but between the easiness of two tests. It will be high if the two tests are easy, low if they are difficult. Its size will be in no way dependent on whether the abilities measured are or are not closely related (*i.e.*, on whether a high score in one test is or is not an indication of a probable high score in the other test).

Adams, however, seems not to realize the importance of this difference. Between Tables 4 and 5 of his part I, he switches over from his results to some of Spearman's without mentioning that Spearman's correlations are of a totally different kind, being correlations between testees and not between orders of test material. This confusion also is essential to Adams's argument, for once it is realized that Tables 5, 6 and 7 are not derived from the same sort of correlations as Adams's own data, it is seen that his statement that the quantity obtained from these latter tables by calculation of $\sqrt{r_{ab} \cdot r_{ac}/r_{bc}}$ is r_{ax} (the correlation of test a with its standard) is without foundation. The quantity is Spearman's r_{ag}, and is actually calculated by means of Spearman's own formula. The smallness of the discrepancy from the theoretically expected value which Adams points out is, of course, exactly the result that would follow from Spearman's theory.

When error arises from confusion between two different things through indicating them by the same symbol, it may best be removed by using different symbols. Reserving the symbol r_{ab} for the correlation between testees in tests a and b, let us use the symbol r_{AB} for Adams's correlation between the orders in which a single subject places the material in tests a and b.* We may also conveniently make a further distinc-

* a and b being tests in which material is arranged in order. It is difficult to see how Adams's reasoning can apply to any other sort of test. There is, it is true, always a "standard" (a correct answer to the test) but in most tests this standard is not an order and cannot therefore be an item in a correlation coefficient.

tion by using the symbol $\bar{r}_{AB}$, for the mean coefficient obtained from a number of subjects. Similarly, we shall use r_{AX} and $\bar{r}_{AX}$ for correlations between the orders obtained in test A and the standard order.

We have seen that $\bar{r}_{AB}$ is not the same kind of measure as r_{ab}. It is not a relationship between the orders of a group of subjects in two tests. It is not, in fact, a relationship between the testees at all. The r_{AX} of a given subject is simply a way of expressing his test score in a particular kind of test. The mean value of r_{AX} will be large if the test is easy for the group tested and small if it is difficult. $\bar{r}_{AX}$ is simply a measure of the easiness of the test. $\bar{r}_{AB}$ is similarly a function of the easiness of tests *a* and *b*, and will be large if they are easy, small if they are difficult. The question of exactly what function this quantity is of the easiness of the two tests is answered by Adams. He shows empirically that $\bar{r}_{AB} = \bar{r}_{AX} \cdot \bar{r}_{BX}$. It is clear that the errors of any individual in the two tests are likely to be uncorrelated if, as was the case in Adams's work, the steps between different items in the same test are of equal difficulty. The above, therefore, seems a sufficiently likely relationship.

Since r_{AX} and r_{BX} are an individual's test scores in tests *a* and *b*, and $\bar{r}_{AX}$, $\bar{r}_{BX}$ are the mean scores for the whole group of subjects, it is easy to express r_{ab} in terms of these values. The relationship is:

$$r_{ab} = \frac{S\{(r_{AX} - \bar{r}_{AX}) \cdot (r_{BX} - \bar{r}_{BX})\}}{\sqrt{S(r_{AX} - \bar{r}_{AX})^2 \cdot S(r_{BX} - \bar{r}_{BX})^2}}.$$

Plainly, it is an error to confuse this index with r_{AB} which is simply a mean score. r_{ab} may be low while $\bar{r}_{AB}$ is high or r_{ab} may be high while $\bar{r}_{AB}$ is low. If $\bar{r}_{AB}$ is high this means that tests A and B are both easy so that each shows a high correlation with its standard; if r_{ab} is high this means that the abilities measured by *a* and *b* are such that a subject who scores well or badly in *a* is likely also to score about as well or badly in test *b*.

Adams's persistent confusion of a correlation between two sets of scores of the same group of individuals, with the mean correlation between two arrangements of test material, is further shown on page 19 where he refers to $\bar{r}_{A_1A_2}$ (the coefficient obtained by two repetitions of the same discrimination test) as the so-called "reliability." It is difficult to believe that any one has ever called this measure the "reliability" of a test although, at first sight, it looks like a reliability measure, particularly when it is written (as it is by Adams) in a form indistinguishable from the true reliability measure—$r_{a_1a_2}$. Its properties, however, are entirely different. The reliability of a test is the extent to which it gives the same order of scores to testees on repetition. Plainly, $\bar{r}_{A_1A_2}$ does not measure this. If, for every subject, $r_{A_1X} = r_{A_2X}$ then the test is completely reliable, but $\bar{r}_{A_1A_2}$ may be large or small according to whether the $\bar{r}_{AX}$ is large or small (successive errors of the same subject being uncorrelated, $\bar{r}_{A_1A_2} = \bar{r}_{A_1X} \cdot \bar{r}_{A_2X}$ or $\bar{r}_{AX}{}^2$).[4] In other words, $\bar{r}_{A_1A_2}$ is also simply a measure of the easiness of the test, not of its reliability. This index may be small when $r_{a_1a_2}$ is large (if the test is difficult but reliable) or *vice versa.*

The correlation coefficients of Adams are, therefore, of a totally different kind from the correlations between testees on which Spearman's theory of general ability is based. The mathematical work of Spearman considers, however, a more general problem of which the two-factor theory of ability is only a special case. Adams's discovery that his coefficients fulfill the mathematical conditions of this general theory is entirely consistent with this part of Spearman's work and is to be expected from it.

The essence of Spearman's mathematical work may be expressed as follows. If there are four variates, a, b, c and d, which are intercorrelated as a result of the correlations of each of them with a fifth variate Z, but have no partial intercorrelation when Z is held constant, then, and then only, will

[4] H. F. Adams, *op. cit.*, p. 19.

certain mathematical relationships hold between the intercorrelations of a, b, c and d. The principal of these relationships are: the *tetrad equation* ($r_{ab} \cdot r_{cd} - r_{ac} \cdot r_{bd} = 0$) and the equations $r_{ab} = r_{aZ} \cdot r_{bZ}$ and $r_{aZ} = \sqrt{r_{ab} r_{ac} / r_{bc}}$. The variate Z may be called the *common factor.*

This system of equations may clearly hold between variates of widely different kinds, and the nature of Z will depend on the nature of a, b, c and d. Adams seems to suppose that Spearman thinks that Z is always "general ability." This is not the case. Z is general ability only when a, b, c and d are the series of scores of a group of experimental subjects in four tests of different abilities. Spearman, himself, in 1904 showed the general mathematical principles involved in the two-factor theory might be illustrated by a relationship in which the series of numbers aimed at in a target played the part of common factor,[5] thus anticipating Adams's demonstration that the general mathematical condition could be fulfilled by a suitable series of terms connected in the required way with the order of a standard. Adams's contribution is that he has demonstrated empirically that the same system of relationships hold for the mean correlations between individual orderings in discrimination tests. They may well also hold between many other kinds of similar quantitative material, as for example, Rugg's rainfall tables quoted by Adams.[6]

In Adams's data, the common factor is, as he correctly points out, the order of the standard material. This is in no sense a contradiction of Spearman's contention that when the correlated data are the series of scores in tests of ability, the common factor is general ability. Adams's error lies in the fact that he ignores the totally different character of the correlated items in the two cases, and thinks therefore that Spearman's general ability must be identical with his own x (the standard order in an arrangement test). As well might

[5] *Amer. Jour. Psych.*, 1904, XV, p. 271, quoted by Adams, JOURNAL OF APPLIED PSYCHOLOGY, 1931, XV, p. 31.

[6] *Op. cit.*, p. 30.

Spearman have said in the above example, "Therefore, general ability is a target." The correct statement is that there is an analogy between the mean orders of individual responses in discrimination test and the orders of different testees' scores in that or any other mental test.

Let us examine the analogy more closely. In Adams's examples, there is a certain order of the discriminatory material—the correct or standard order. In two different tests the material is placed by the subject in an order which is in each case a shot at the correct order but not a completely successful one. The correlations with the correct order are r_{AX} and r_{BX}. The deviations from the correct order of the material in these two tests are uncorrelated with one another; therefore, the correlation which exists between the two orderings of the material is entirely dependent on the extent to which each approximates to the standard order, so the relationship $\tilde{r}_{AB} = \tilde{r}_{AX} \cdot \tilde{r}_{BX}$ holds.

In Spearman's work we are dealing with the orders or scores of different subjects instead of orders of test material for the same subject. Otherwise the situation is similar. There is a certain possible order of the subjects—that in which they would be arranged in order of general intelligence. In each of two different tests, these subjects are arranged in an order which is a more or less close approximation to the order of their general ability but which is not exactly the order of their general ability. The correlations with the order of their general ability are r_{ag} and r_{bg}. The deviations of the orders in the two tests from the order of the subjects' general ability (due to the different specific abilities required by the two tests) are uncorrelated with one another. Therefore, the correlation which exists between the orders of the subjects in the two tests is entirely dependent on the extent to which each of these orders approximates to the order of the subjects in general ability, so the relationship $r_{ab} = r_{ag} \cdot r_{bg}$ holds.

It is clear that there is an analogy between the part played by the standard in correlations between orderings of discriminatory test material and that played by general ability in correlations between the orders or scores of subjects in different tests. The analogy is shown by identity of mathematical relationships between correlation coefficients. The common factor producing this identity of mathematical relationship is, in the first case, the standard order of the test material; in the second case, it is the order of the testees which would have been obtained if they could have been arranged by some method measuring general ability alone.

It follows that whether or not Adams's method of distinguishing between subjective and objective scales proves to be a valid and useful one, his results have no bearing whatever on the theory of general ability. It is not necessary to examine also his observations on specific factors and group factors since these rest on the same fundamental fallacy.

THE INTELLIGENCE AND DRAWING ABILITY OF YOUNG MEXICAN CHILDREN

H. T. MANUEL
University of Texas
AND
LOIS S. HUGHES
Brownsville Junior College

In 1928 Mrs. Hughes made a comparative study of the intelligence of Mexican and non-Mexican children.[1] The study was based upon the Goodenough Intelligence Test, which requires only the drawing of a man. Drawings were selected by a process of sampling from about twelve thousand made by the white children in the first four grades of the San Antonio Public Schools, 440 for the Mexican and 396 for the non-Mexican children. In most cases the child, if he could write, classified himself as Mexican or non-Mexican. In other cases the classification was made by the teacher or the investigator.

In the Goodenough test an index of a child's mental development is secured from his drawing of a man. This is done by an ingenious method of scoring based largely upon the number of parts of the figure represented by the child. The drawing itself, however, is objective and may be studied from the point of view of ability in drawing. There is nothing in the directions to indicate the use that is to be made of it.

Accordingly, the drawings previously used in the study of intelligence were examined from the standpoint of their quality as drawings. They were graded by a supervisor of drawing and a former teacher of drawing on the basis of the Thorndike Scale.[2] The drawings, then, have been the basis

[1] Unpublished thesis in the library of the University of Texas.

[2] Grateful acknowledgment is made to Miss Thelma Whaley and Mrs. Wright Stubbs for grading these drawings. The intelligence ratings are in the main those of Mrs. Hughes, though a few were slightly re-

of both intelligence ratings and drawing scores. In the following paragraphs data will be presented to show (1) the correlation between intelligence and drawing and (2) a comparison of the relative abilities of Mexican and non-Mexican children.

TABLE I

*Correlation of Thorndike and Goodenough Scores**

By Ages

AGE	COEFFICIENT	
	Mexican	Non-Mexican
7	.83 ± .04	.72 ± .04
8	.78 ± .03	.69 ± .04
9	.72 ± .04	.74 ± .03
10	.86 ± .02	.78 ± .04
11	.82 ± .03	
12	.65 ± .05	
Median r	.78	.73

By Grades

GRADE	COEFFICIENT	
	Mexican	Non-Mexican
1	.79 ± .02	.75 ± .03
2	.68 ± .04	.81 ± .03
3	.65 ± .04	.63 ± .04
4	.63 ± .06	.72 ± .03
Median r	.67	.74

* Coefficients are uncorrected.

CORRELATION OF INTELLIGENCE AND DRAWING

Coefficients of correlation for each of the four grades and for the ages which are best represented numerically in these grades are given in Table I.

vised by Miss Myrtle Killam. Lee Davenport contributed to the statistical work. Miss Elma A. Neal, Assistant Superintendent of the San Antonio Schools, kindly made it possible to secure the original drawings. The University of Texas Fund for Research in the Social Sciences gave financial assistance in a part of the study.

The coefficients of correlation between the Thorndike (drawing) and the Goodenough (intelligence) scores are remarkably high, ranging from .63 to .86. Apparently the elements which make a child's drawing of a man a good index of mental ability are much the same as, or are commonly found with, the elements which make it a good drawing. It will be observed that the coefficients of correlation are less in the third and fourth grades than in the first and second grades.

COMPARISON OF THE MEXICAN AND NON-MEXICAN CHILDREN

From Table II it is evident that the Mexican children are retarded greatly, being from a year and four months to two years and four months older grade for grade. It may be seen too that their ages are much more variable in each grade. The piling up of Mexican children in the first grade is shown by the relative numbers in the four grades.

TABLE II

Ages of Children*

	NUMBERS		MEDIAN AGE		INTERQUARTILE RANGE OF AGES	
	Mex.	Non-Mex.	Mex.	Non-Mex.	Mex.	Non-Mex.
Grade 1	196	115	8–11	7–7	1–10	0–9
Grade 2	102	96	10– 9	8–9	2– 2	1–1
Grade 3	89	92	11–11	9–7	2– 7	1–0
Grade 4	53	93	12– 6	10–9	2– 2	1–8

* Ages are in years and months. The legal age of entrance to the first grade at the time of this study was seven years.

A comparison of the scores of Mexican and non-Mexican children by school grades and by ages is given in Tables III and IV, respectively.

In the first grade the Mexican children are slightly below the other children in both intelligence and drawing, but as they advance in grade they gain a substantial lead. In the intelligence test in which age norms are available the superi-

TABLE III

Goodenough Intelligence and Thorndike Drawing Scores

By Grades

GRADE	MEAN SCORE				SIGMA OF DISTRIBUTION			
	DRAWING		INTELLIGENCE		DRAWING		INTELLIGENCE	
	Mex.	Non-Mex.	Mex.	Non-Mex.	Mex.	Non-Mex.	Mex.	Non-Mex.
1	6.3	6.9	20.0	20.0	1.5	1.4	7.4	5.7
2	7.3	7.4	27.2	26.1	1.4	1.3	7.6	7.4
3	8.5	7.8	31.6	28.3	1.9	1.2	8.6	6.5
4	9.2	8.4	35.4	32.5	2.2	1.2	8.2	7.3

TABLE IV

Goodenough Intelligence and Thorndike Drawing Scores

By Ages

AGE	MEAN SCORE				SIGMA OF DISTRIBUTION			
	DRAWING		INTELLIGENCE		DRAWING		INTELLIGENCE	
	Mex.	Non-Mex.	Mex.	Non-Mex.	Mex.	Non-Mex.	Mex.	Non-Mex.
7	5.4	7.0	16.4	21.0	1.5	1.4	5.7	5.4
8	6.3	7.1	20.0	24.2	1.6	1.2	6.7	5.7
9	7.2	7.8	22.5	28.8	1.3	1.2	7.5	7.3
10	7.6	8.2	26.7	30.1	1.7	1.4	9.8	7.8
11	7.2		28.5		1.6		7.4	
12	8.4		31.5		2.4		8.9	

ority of the Mexican children in the fourth grade is equivalent to about nine months of mental age. The Mexican children are more variable in every grade.

Because of the lack of comparable norms it is hard to evaluate the standing on the Thorndike test. Childs reports a gain of 3.5 to 3.97 points from Grade Ia to Grade VIa, but his results are not based on the same subject as was used in our work.[3] Goodenough mental ages and intelligence quotients are shown in Table V.

TABLE V

*Mental Ages and IQ's**

	MENTAL AGES		INTELLIGENCE QUOTIENTS	
AGE	Mexican	Non-Mexican	Mexican	Non-Mexican
7	7–1	8–3	94	110
8	8–0	9–1	94	107
9	8–7	10–2	90	107
10	9–8	10–6	92	100

* The intelligence quotients shown here are found simply by dividing the mean mental age by the mean age of each age group.

When the comparison is made by ages, the Mexican children have lower average scores and a greater variability at every age level. The average difference in mean intelligence scores of ages seven to ten, the ages for which valid comparisons are available, is equivalent to a little more than a year of mental age. It will be noted, however, that the difference in scores between the seven year olds and the ten year olds is greater for the Mexican children. It is possible, of course, that some of the more able non-Mexicans of ten years of age were in higher grades and were therefore not included in the test. The immaturity of the Mexican children in drawing at age seven is especially evident.

[3] Childs, H. G. "Measurement of the Drawing Ability of 2177 Children in Indiana School Systems by a Supplemented Thorndike Scale." *Jr. Ed. Psy.* 6: 391–408, 1915.

CONCLUSIONS

A study of the drawings from the standpoint of intelligence and of ability in drawing shows these traits to be closely related in so far as they are measured by the technique adopted. This relation decreases, however, with advance in school grade. It is suggested that probably the factors which differentiate talent in drawing are less obvious in the first grade than they are in later grades.

An examination of the age-grade location of the Mexican children included in this study shows that they are seriously retarded. Other more extensive studies in the Southwest show that this condition is a common one among Mexican children.[4]

In general the average ability of the Mexican children both in intelligence and in drawing compares favorably, *grade for grade,* with that of the other children. In grades three and four (in intelligence, grade two also) the Mexican children have the higher means. So far then as the absolute level of these abilities is concerned, the Mexican children should be able to compete successfully with other children of their school grades. In fact, their advancement through the grades seems to lag behind their growth in mental ability.

A comparison *by ages,* however, is less favorable. At each age level (7–10) the Mexican children have lower scores both in drawing and in intelligence. In intelligence the difference is equivalent to about a year of mental age.

The lower standing of the Mexican children in intelligence agrees with the findings of various investigators. The belief among many observers, however, that Mexican children are gifted in drawing and handwork is not supported by a comparison of the scores at any age level. It may be that the apparent talent of Mexican children in drawing is largely a matter of training and interest or else that the test used in this study is not a suitable one to reveal their ability.

[4] See Manuel, H. T. *The Education of Mexican and Spanish-Speaking Children in Texas.* 1930.

STUDIES OF TYPOGRAPHICAL FACTORS INFLUENCING SPEED OF READING

VIII. Space Between Lines or Leading*

DONALD G. PATERSON and MILES A. TINKER
University of Minnesota

Although space between lines (leading) has been recognized by all authorities on typography as a factor of major importance in legibility, there is a surprising absence of detailed and definite evidence available to guide hygienic printing practice. For this reason, leading becomes an important factor to be considered in our general program of research on typographical factors influencing speed of reading.

REVIEW OF LITERATURE

A survey of available literature reveals only a few experimental studies of the problem before us. These fall into two classes: (a) investigations of leading in book and magazine typography using adults or children as subjects, and (b) studies of leading under more special conditions such as telephone directories and newspaper typography.

The earlier authorities tended to base recommendations on common sense observations and clinical practice. For example, Cohn[1] emphasized the importance of leading, but Javal[2]

* The writers are grateful to the following for aid in conducting this study: Educational Test Bureau, Minneapolis, for permission to reprint the Chapman-Cook Speed of Reading Tests, Forms A and B, for experimental purposes; the University of Minnesota Graduate School, for research grant; and the instructors in charge of Freshman English classes in the College of Engineering, the University of Minnesota, for permission to test their students.

[1] H. Cohn. *The hygiene of the eye in schools.* London, 1886.

[2] E. Javal. *Physiologie de la lecture et de l'écriture.* Paris, Feliz Alcan, 1906.

stated that the absence of leading does not reduce legibility and is therefore a luxury. Huey,[3] basing his recommendations on the statements of his predecessors, minimized the importance of leading and emphasized size of type as a much more important factor.

Griffing and Franz[4] seem to have been the first to put the matter to an experimental test. Using text material in 5 point type (Pearl) presented to a small number of adult subjects, they found a slight increase in legibility, as measured by the illumination threshold method, when leading was increased from .8 mm. to 1.3 mm. This result, however, is limited in practical significance primarily because of the extremely small size of type employed.

Bentley[5] conducted an extensive series of experiments to determine the effect of different amounts of leading on legibility. Using 12 point monotype closely approximating "news Gothic," in 20 pica line widths, Bentley measured speed of oral reading a ten-line passage at a given distance from the eye when the text was set solid, and when the text was leaded from 1 to 9 points by steps of one. However, results are reported only for 0, 3, 6, 7, 8, and 9 point leading because "the others appeared to have no special significance." Therefore, estimates regarding the effect of 1, 2, 4, and 5 point leading must be inferred from the curves presented. In general, Bentley found that every amount of leading except 9 point facilitates speed of reading aloud. Table 1 gives the average time for reading the ten-line passages leaded as indicated above. These figures are taken from the top curve in Graph III (p. 50) showing the averages for all observations at all distances. In addition to the data taken directly from Bentley's report, we have included in the third column of the table

[3] E. B. Huey. *The psychology and pedagogy of reading.* New York, Macmillan, 1909.

[4] H. Griffing and S. I. Franz. *On the conditions of fatigue in reading. Psychol. Rev.*, 1896, 3, 513–530.

[5] M. Bentley. *Leading and legibility. Psychol. Monog.*, 1921, 30, No. 136, 48–61.

TABLE 1

Summary of Bentley's experiment concerning effect of leading on speed of oral reading (Monotype, 12 point, 20 pica line width)

LEADING	AVERAGE TIME IN SEC.	PERCENTAGE INCREASE
0	24.0	0.0
3	22.7	+ 5.4
6	22.6	+ 5.8
7	21.3	+11.3
8	23.0	+ 4.2
9	24.2	− 0.8

the difference in per cent between the reading time for material set solid and each degree of leading. It is apparent that amounts of leading from 3 to 8 point facilitate reading whereas 9 point leading shows a slight retarding effect. The increase in reading time for 7 point leading is striking. It is to be noted, however, that 3, 6, and 8 point leading yield an appreciable increase of about 5 per cent. Should the findings with respect to 7 point leading be verified by other experimenters, it is obvious that the advertiser could profitably apply the principle although it is doubtful whether it would be feasible in general printing practice.

Bentley also obtained similar data for 6 point type and 9 point type. These type sizes were approximated through photographic reductions of the original 12 point text. Although the general shape of the curves remains the same for these two reductions, it is doubtful whether the obtained curves can be taken as representative for these type sizes. The difficulty arises from the fact that photographic reductions do not duplicate printing conditions since printing in 6 point and 9 point type as compared with 12 point type does not involve a proportional change in height and width of letters and spaces, such as occurs in the photographic process. In spite of this criticism it should be recognized that Bentley has opened up an important problem with respect to a possi-

ble constant ratio between interlinear spacing and size of type which will produce optimal reading conditions.

Very little work has been done on the problem of optimal leading for reading material in the lower grades. Blackhurst,[6] on the basis of actual tests of school children in the first four grades, concluded that as much as 10 point leading might be desirable in the first grade, but that there is no advantage of leading greater than 1 point in the second, third or fourth grades. This conclusion, however, is greatly limited in significance since size of type and line width were not specified.

Hovde,[7] who limited his studies to newspaper typography, found no consistent effect of leading on speed of reading. Type sizes varied between 6½ and 9 point and leading varied by quarter points from material set solid to 1 point. Line width was constant at 12½ picas. It is possible that Hovde's inconclusive results reflect merely the restricted range of leading differences employed.

Finally we may mention the work of Baird[8] and of Lyon[9] who investigated the speed and accuracy with which telephone directories can be consulted when the names and numbers are set solid and when one point leading is inserted. Baird discovered that 1 point leading increased efficiency in the use of the telephone book while Lyon found that 0 and 1 point leading yielded approximately equal results. But these investigations throw no light on our present problem since the authors were concerned with such a specialized kind of printing.

It is apparent from this survey of the literature that present knowledge concerning the effect of leading is fragmentary

[6] J. H. Blackhurst. *Investigations in the hygiene of reading.* Baltimore, Warwick and York, 1927.

[7] H. T. Hovde. *The relative effects of size of type, leading and context.* J. APPL. PSYCHOL., 1929, 13, 600–629, 1930, 14, 63–73.

[8] J. W. Baird. *The legibility of a telephone directory.* J. APPL. PSYCHOL., 1917, 1, 30–37.

[9] O. C. Lyon. *The telephone directory. Bell Tel. Quart.*, 1924, 175–185.

and inconclusive. There is urgent need for additional experiments which will duplicate normal reading conditions and allow definite measurement of the effects on legibility of various amounts of leading with all typographical conditions clearly specified.

PROCEDURE

A. Preliminary Experiment

In the present study, the procedure used differs from our previous studies in only one respect. In the earlier experiments we used the Chapman-Cook Tests as originally printed, *i.e.*, 30 paragraphs of 30 words each. This arrangement results in frequent breaks in the continuity of printing and consequently an unusual amount of white space occurs between every third and fourth line. In a sense this amounts to the insertion of an unusual amount of leading between every third and fourth line, a condition which is undesirable for an investigation aimed directly at the influence of leading itself. To counteract this defect (due to shortness of the Chapman-Cook paragraphs) it was decided to rearrange the paragraph spacing by combining each set of five paragraphs into a single printing unit. The effect of this was to give the appearance of a printed page of six paragraphs (3 to a column). Thus, when leading was introduced, there was greater opportunity to measure this effect as a single variable.

As mentioned in our published reports on typography, the validity of our whole program is dependent upon equivalence in difficulty of Forms A and B. This has been demonstrated for the usual Chapman-Cook Test arrangement. However, we know that such a factor as previous exposure to the tests destroys equivalence.[10] For this reason it might be argued that a rearrangement of the 30 paragraphs into six printing units might upset equivalence. Therefore, as a preliminary

[10] D. G. Paterson and M. A. Tinker. *Studies of typographical factors influencing speed of reading: IV. Effect of practice on equivalence of test forms.* J. Appl. Psychol., 1930, 14, 211–217.

step in our study of leading it was thought advisable to determine equivalence under this new printing arrangement. To determine this we gave the tests to summer school students in psychology at the University of Minnesota, using two arrangements: (a) control group of 100 students to whom the tests were given in the usual printing arrangement (10 point set solid, 19 pica line width, Scotch Roman on egg-shell paper stock) and (b) experimental group of 94 students to whom the tests, arranged in six printing units, were given. All other typographical factors were identical with those in the control group. We use the terms experimental and control groups for convenience only since the thirty paragraph arrangement and the six printing unit arrangement were systematically distributed in each class tested. Results are given in Table 2.

TABLE 2

Relative difficulty of Forms A and B when printed in the usual 30 paragraph arrangement and when printed in a six printing unit arrangement

TEST GROUP	TEST FORM	MEAN	$S.D._M$
Control group ($N = 100$)	A	19.09	.48
	B	18.68	.47
Experimental group ($N = 94$)	A	17.78	.53
	B	17.91	.47

Table 2 shows that Forms A and B are practically equivalent for both printing arrangements. Such differences as exist between Forms A and B (.41 for control group and .13 for experimental group) may well be attributed to chance since the differences between the means are smaller than the standard deviations of the means themselves. These results indicate that the rearrangement of paragraphing does not destroy equivalence of Forms A and B, although the new arrangement does slow up rate of reading to a considerable degree. The fact that the six unit printing arrangement is

read more slowly is to be expected since it violates orthodox paragraphing principles. This disturbance is distributed equally in both test forms, and these forms, therefore, may be used to measure the influence of any typographical variation introduced into Form B as compared with Form A.

B. Experiment Proper

In our investigation of the influence of leading we compared efficiency in reading Form A of the Chapman-Cook Speed of Reading Test (set up in six printing units, 10 point Scotch Roman, 19 pica line width, set solid on egg-shell paper stock) as a standard, with efficiency in reading Form B of the Chapman-Cook Test, identical typographically with Form A but differing in leading as follows: set solid, 1 point leading, 2 point leading, and 4 point leading. Variations in leading beyond 4 points were not included in our study because they would fall outside the range of normal printing practice which rarely goes beyond 2 point leading.

It must be understood of course that the present study is concerned with adult reading only. Were we to investigate the influence of leading on speed of reading at the child level (grades I to IV) it would be important to extend the range of leading employed beyond 4 points.

The following four test groups of 100 college students each were used: Group I was introduced as a control group to demonstrate again that Forms A and B are approximately equal under the conditions of this experiment. This group serves as a further check on the control group in Table 2.

Group I. Form A, set solid, Form B, set solid
Group II. Form A, set solid, Form B, 1 point leading
Group III. Form A, set solid, Form B, 2 point leading
Group IV. Form A, set solid, Form B, 4 point leading

In all, 400 students in Engineering Freshman English classes at the University of Minnesota were tested. None of these students had previously been exposed to the Chapman-

Cook Test. The four sets of tests used were systematically distributed for administration in each of the classes.

The testing was conducted in accordance with our previous practice as regards testing procedure and time limits.[11]

RESULTS

Table 3 shows that 1 point leading does not facilitate speed of reading as compared with text set solid. However, 2 point and 4 point leading definitely increase speed of reading. In Group III, for example, the third column shows that on the average 15.84 paragraphs of Form B, 2 point leading, was read within the time-limit, whereas only 14.74 paragraphs of Form A, set solid, was read within the time-limit. This is a difference of 1.10 paragraphs as shown in column 6, and indicates that Form B, 2 point leading, is read 7.5 per cent faster than material set solid. The statistical significance of this difference is shown in column 10. It is apparent that a statistically certain difference appears only when leading is 2 point or 4 point, with a slight advantage in favor of 2 point.

On the basis of these results one would not be justified in drawing the conclusion that 2 and 4 point leading will necessarily increase speed of reading in all printing situations. It must be remembered that the favorable effect of 2 and 4 point leading shown in this experiment may turn out to be true only for material set in 10 point type with 19 pica line width. Further light on the extent to which 2 and 4 point leading will exert a favorable influence on reading rate must await similar experiments with additional type sizes and line widths.

Judging from Bentley's experiment, which showed that the effectiveness of relative proportions of leading remained approximately the same for 12 point, 9 point, and 6 point type, it is probable that optimal leading will vary somewhat with type size, being smaller for small type and larger for large type. It would be premature, however, to make such a gen-

[11] D. G. Paterson and M. A. Tinker. *Studies of typographical factors influencing speed of reading: II. Size of Type.* J. APPL. PSYCHOL., 1929, 13, 120–130.

TABLE 3

Influence of leading

Differences are for the mean score on Form A, 10 point, 19 pica line width, Scotch Roman, egg-shell paper stock, set solid minus the mean score on Form B identical typographically with Form A but varying in leading from 0 to 4 points (0, 1, 2, and 4 point leading). The mean score is the average number of paragraphs of 30 words each in the Chapman-Cook Speed of Reading Test (printed in a six unit arrangement) read in $1\frac{1}{4}$ minutes. In each sub-group, N = 100 freshmen in College of Engineering English classes, University of Minnesota.

TEST GROUP	TEST FORM AND LEADING	MEAN	$P.E._{dist.}$	$P.E._{M}$	DIFFERENCE BETWEEN MEANS IN			$P.E._{diff.}$*	$\frac{D}{P.E._{diff.}}$
					Paragraphs	Words per sec.	Per cent		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
I	A, 0 point	15.28	2.83	.28	.34	0.09	2.2	.15	2.24
	B, 0 point	15.62	2.58	.26					
II	A, 0 point	15.70	2.89	.29	.19	0.05	1.3	.17	1.15
	B, 1 point	15.89	2.40	.24					
III	A, 0 point	14.74	2.79	.28	1.10	0.31	7.5	.15	7.26
	B, 2 point	15.84	2.60	.26					
IV	A, 0 point	15.63	2.77	.28	.78	0.22	5.0	.16	4.81
	B, 4 point	16.41	2.52	.25					

* $P.E._{diff.}$ values in column 9 were computed on basis of formula for correlated measures: see H. E. Garrett, Statistics in Psychology and Education, Longmans, Green & Co., N. Y., 1926, pp. 286–288. The correlations for the four sub-groups were: Group I, $+.85 \pm .02$; Group II, $+.82 \pm .02$; Group III, $+.84 \pm .02$; Group IV, $+.82 \pm .02$.

eralization since Bentley's results for 6 point and 9 point text were based on photographic reductions of 12 point type. As stated before, such reductions do not duplicate printing conditions.

Our findings throw doubt on the wisdom of current printing practice which frequently employs 1 point leading. They also demonstrate the desirability of a more general use of 2 point leading with 10 point type and approximately 19 pica line widths. Similarly, the extra space required for 4 point leading does not yield as satisfactory results as can be obtained with 2 point leading.

SUMMARY AND CONCLUSIONS

1. The present study was undertaken to discover whether 1, 2, and 4 point leading would exert a favorable effect on speed of reading in comparison with material set solid. Since our earlier studies indicated that 19 pica line width was optimal for 10 point type when set solid, we employed these printing specifications for our present study of leading.

2. A special arrangement of the Chapman-Cook Speed of Reading Test was employed and demonstrated to yield equivalent scores on Forms A and B. Thus varying amounts of leading could be inserted in Form B, and scores on B compared with performance on Form A, set solid. In all, 400 college students were used in this experiment, none of whom had previously taken the tests.

3. One point leading does not facilitate speed of reading as compared with text set solid.

4. Text material printed with 2 point leading is read 7.5 per cent faster than material set solid.

5. Text material printed with 4 point leading is read 5.0 per cent faster than text set solid.

6. The advantage of 2 point leading may hold only for the particular size of type and width of line used in this investigation. Further studies in this series will investigate the effect of leading for additional type sizes and line widths.

THE RATING OF TEMPERAMENTAL QUALITIES

K. H. R. EDWARDS
Psychological Laboratory, Cambridge, England

1. Aim of Paper.
2. Method.
3. Analysis of Results.
4. Conclusions.

I. *Aim of Paper.*

According to L. Klages, "Temperament obviously is a single disposition which can be graduated indefinitely."[1] This definition implies that, since temperament reveals itself through particular temperamental qualities or traits, these, too, can be graduated. But the most common means of estimating or rather indicating the latter, the Rating Method, is widely criticised in that it presupposes unity of the traits in question and in this way stresses their relative independence rather than their continuity. The aim of this paper is to show that contrary to this opinion the Rating Method used in a particular way does show Temperamental qualities and we may infer Temperament to be capable of infinite graduation as Klages has indicated.

II. *Method.*

Among the ways of indicating or assessing temperamental qualities are Performance Tests, Handwriting Tests and Emotional Tests. It is becoming more widely recognized that these tests should not represent to be tests of Temperament in the same way we speak of tests of Intelligence, but rather they should be tests through which Temperamental qualities are indicated.[2] Each test is usually designed to indicate the

[1] L. Klages. The Science of Character, London, 1929, page 149.

[2] F. C. Bartlett. Temperament and Social Status. J. N. I. I. P. London, 1927. 111. 8, 401–405.

presence or absence of some particular trait or group of traits. In this way certain traits receive undue prominence and attention whilst others are relatively neglected. Owing to this disparity graduation of traits is almost impossible.

An objection levelled against Performance and Handwriting Tests is that no matter how great their number and apart from any statistical reliability or unreliability which they may reveal they cannot possibly indicate the whole range of temperamental qualities, but only those qualities which find expression through motor mechanisms such as recklessness, deliberateness, nervousness and constructiveness. They cannot normally indicate traits such as boastfulness, sociability, cooperativeness and imitativeness which appear in the social group. To assess these direct observation and rating appears the only satisfactory way.

Thus in order to secure comprehensiveness and evenness of assessment the rating of an extremely wide field of temperamental characteristics was attempted. For the wider the field the better is the chance of infinite graduation. As a necessary preliminary a comprehensive list of terms used to describe temperamental qualities or traits was compiled from a multitude of sources. Complementary terms such as Optimism, Pessimism, and different terms describing the same common quality, but yet possessing slight differences, such as Sociability and Social Participation, were included.

The final list included 70 terms which were arranged alphabetically. A class of 20 students of psychology (who were thus well acquainted with the terms employed and their particular significance for this purpose) were then asked to rate their tutor, whom they had known for 18 months both inside and outside the lecture room, in the qualities enumerated. A six point scale 0, 1, 2, 3, 4, and 5 was used, 0 indicating the complete absence, and 5 possession, of the trait concerned. The traits were read out by the tutor himself, ample time being given for each rating. At the conclusion the papers were collected and the ratings totalled. The traits were then

rearranged in numerical order of scoring. The list is given below:

Total No. of marks awarded	*Trait*	Total No. of marks awarded	*Trait*	Total No. of marks awarded	*Trait*
95.	Concentration	75.	Dynamic drive	46.	Nervousness
94.	Ambition	75.	Calmness under pressure	44.	Moodiness
92.	Attention to detail	75.	Optimism	42.	Imitativeness
91.	Reliability	75.	Self-estimation	39.	Suggestibility
89.	Confidence	74.	Initiative	38.	Excitability
89.	Control of speech	73.	Pushfulness	34.	Changeability
87.	Stability	72.	Co-operativeness	31.	Melancholic tendencies
87.	Self-control	71.	Sociability	31.	Recklessness
87.	Self-confidence	68.	Social participation	29.	Emotional instability
87.	Deliberateness	68.	Patience	28.	Impulsiveness
86.	Activeness	67.	Daring	28.	Overemotionality
85.	Thoughtfulness	66.	Assertiveness	28.	Vanity
84.	Quickness	65.	Inhibitive tendencies	27.	Submission
83.	Persistence	61.	Contentedness	24.	Inferiority feelings
83.	Accuracy & speed of movement	59.	Buoyantness	24.	Pessimism
82.	Caution	59.	Conservativeness	23.	Timidity
82.	Energy	58.	Superiority traits	18.	Irresolution
82.	Seriousness	53.	Domination	17.	Boastfulness
81.	Foresight	50.	Extroversion tendencies	16.	Lack of planning ability
78.	Emotional balance	48.	Introversion tendencies	14.	Obtrusiveness
78.	Fearlessness	48.	Impatience	10.	Instability of interest
78.	Resourcefulness	47.	Restlessness	8.	Slowness
78.	Plodding	46.	Aggression	4.	Sluggishness
75.	Hope				

III. *Analysis of Results.*

1. The group rating shows the traits to be grouped in an almost continuous arithmetical range. Ratings of other individuals confirmed this tendency.

2. There appears a rational arrangement among the ratings. Opposite traits are found at opposite ends of the scale (*e.g.*) activeness is rated 86, sluggishness 4; concentration is rated 95, instability of interest 10; stability is rated 87 and irresolution 18. Also, where one trait is indicated by more than one term, these are found grouped together (*e.g.*) co-operativeness 72, sociability 71, social participation 68. Superiority and domination appear with 58 and 53 respectively. Impatience and restlessness are adjacent as are also inferiority feelings and timidity.

3. The positions of these traits form an internal check, as it were, on the validity of the ratings, which as one can judge from the foregoing and the rating of extroversion and introversion as 50 and 48 respectively, is quite high.

4. The method does not attempt to measure temperamental qualities, but to indicate their relative importance, which is in accordance with an accepted aim of temperamental test methods.

5. Graphically, making allowance for the inclusion of duplicated terms in some instances, the individual temperament would be represented by a straight line inclined to the horizontal axis and intersecting the vertical axis at a point representative of the highest rated trait, the intersection of the horizontal axis being proportional to the lowest rated trait. Individual temperament graphs would differ according to the gradient.

IV. *Conclusions.*

1. By means of a group rating method we have shown that temperamental qualities may be graduated in an arithmetical series which is in accordance with the definition of Temperament given by Klages. The distribution of traits will vary from individual to individual as will the sum total of the ratings but, in each case, graduation occurs. Further, the practice of self-rating is obviated together with its personal factor.

2. The method is capable of further application and experiment. By means of the latter the most suitable lists of traits and the most suitable experimental conditions will be evolved.

3. On this basis of continuity and graduation of traits, the widely fluctuating profiles used to illustrate certain test methods now appear incomplete. They are consequent upon the selection of too few traits or the rating by too few persons. In the former case, the traits are given a unitary significance which disappears when a wider range of qualities and a larger number of ratings are taken.

4. The majority of the so-called temperament tests as ordinarily used do not measure or indicate the individual temperament but temperamental differences from individual to individual. The group rating method, however, involves the full consideration of the sum total individual characteristics. The method will always be available, irrespective of progress of other forms of temperament estimation, as a means of direct check on the validity of the latter.

MEASUREMENT OF THE APPEAL OF PERFORMANCES IN THE THEATER

RICHARD LEDGERWOOD
University of Maine

It was desired to determine the facility with which comparative esthetic judgments of such elaborate and involved sensory patterns as are provided by the units of a variety performance on the stage can be made. For the purpose of measuring the appeal of theatrical performances, the methods of measurement by relative position are of course the most readily available. Both the order-of-merit and the scale-of-values method would seem to be suitable. The specific object of the investigation to be reported was to try out the order-of-merit method.

A performance in Saint Louis of Nikita Balieff's Théâtre de la Chauve Souris, which the writer and a number of his acquaintances attended in November, 1929, supplied the opportunity. This presentation consisted of a musical overture and nineteen divertissements ranging from travesty and farce comedy to interpretative sketches and scenes from opera.

The reagents were 17 graduate students and faculty members, 8 men and 9 women participating. The procedure was simple. The names by which the 19 acts went in the printed program were set down on as many cards. These were presented to the subjects taken individually and the nature of the scenes was briefly reviewed so that there might be no doubt concerning identification of them. The subjects were asked to recall them as vividly as possible and rank them in the order of their general pleasingness. This order was noted down, and the ranks assigned by the separate subjects to the different scenes were afterwards averaged.

The scenes in the order of their presentation on the stage were as follows:

1. Russian Matrimonial Rites
2. The Night Idyl
3. La Laitière de Trianon
4. Harvest Scene
5. The Knife Grinder
6. The Midnight Review
7. Quarrel of Two Gossipers
8. The Minuet
9. Parade of the Wooden Soldiers
10. Les Amours de Jean-Pierre
11. The Russian Cossacks
12. Maxims of a Doorman at Maxime's
13. Ei Ukhnem (The Volga Boatsong)
14. Russian Folk Rhymes
15. Fragment of an Etruscan Vase
16. A Quartet of Merry Artists
17. Street Singers
18. The Celebrated Popoff's Porcelain
19. Chorus of the Brothers Zaitzeff

In the table following they appear in the order of their mean ranks. The mean rank of each scene is presented in the first column. The variability of the means expressed in terms of the standard error is indicated in the second column. These coefficients are interpreted in column three as the range, in rank points, within which the chances are 95 out of 100 that the true mean lies. A notion of the reliability of the means can thus be gained.

The variability of the means is fairly regular throughout the series, the standard error of the mean rank amounting on the average to 1.038. The means are not as reliable as could be desired. They are in general precise only to within about 2 rank points. One exception is the item in first place, which has a precision of half a rank. There is no doubt regarding its premier position as indicated by the composite judgment of 17 persons. It stands 3 ranks ahead of its nearest competitor, a gap larger than any occurring elsewhere in the

THE SCENES	MEAN RANK	VARIABILITY OF THE MEAN	RANGE OF ERROR
Ei Ukhnem (The Volga Boatsong)	2.5	0.27	0.5
The Celebrated Popoff's Porcelain	5.4	1.16	2.3
Les Amours de Jean-Pierre	6.2	1.08	2.1
The Midnight Review	6.3	0.68	1.4
Parade of the Wooden Soldiers	6.6	0.96	1.9
The Knife Grinder	8.1	1.13	2.3
Quarrel of Two Gossipers	9.2	1.04	2.0
Street Singers	9.6	0.82	1.6
La Laitière de Trianon	10.4	1.28	2.6
Fragment of an Etruscan Vase	10.6	0.99	2.0
The Russian Cossacks	11.4	1.07	2.1
The Minuet	12.1	1.46	2.9
A Quartet of Merry Artists	12.2	1.15	2.3
Chorus of the Brothers Zaitzeff	12.4	1.15	2.3
Russian Matrimonial Rites	12.5	1.16	2.3
The Night Idyl	12.7	1.04	2.1
Maxims of a Doorman at Maxime's	12.9	1.02	2.0
Russian Folk Rhymes	14.1	1.04	2.1
Harvest Scene	14.2	1.23	2.5

series. A tendency of the mean judgments to cluster may be noted at several points.

CONCLUSIONS

1. Comparative esthetic judgment of the complex sensory patterns presented by separate acts of a variety performance on the stage has been found practicable, at least in so far as the subjects experienced no difficulty in making responses of the character required.

2. The applicability of the order-of-merit method to such material has been shown.

3. The composite judgments yielded a fair differentiation of the units judged, but so small a number of judges as 17 does not make possible a satisfactorily reliable determination of tendencies.

CONSTRUCTING A PREDICTION CHART (CHARTING LINEAR REGRESSION EQUATIONS)

HAROLD D. GRIFFIN

State Teachers College, Wayne, Neb.

Even when a linear regression equation of the general form

$$\overline{X}_0 = b_{01.2..n}X_1 + b_{02.1..n}X_2 + \cdots + b_{0n.1..(n-1)}X_n + K$$

has been determined, the task of the investigator who desires to use the equation for prediction is by no means completed.

For example, suppose that we have obtained a regression equation, $\overline{X}_0 = 0.06X_1 + 0.43X_2 - 0.87X_3 - 7.35$, from the complete data suggested by Table 1.

Let us predict the score on the criterion for subject 1 from this regression equation. Substituting his scores on tests 1, 2, and 3, which are symbolized as X_1, X_2, and X_3 in the above equation, we have

$$\overline{X}_0 = 0.06 \times 247 + 0.43 \times 237 - 0.87 \times 52 - 7.35$$
$$= 14.82 + 101.91 - 45.24 - 7.35$$
$$\overline{X}_0 = 64.14, \text{ or, briefly, } 64.$$

(The actual score made by subject 1 on the criterion was 61, as can be seen from Table 1.)

If this mathematical process had to be undergone in order to determine the probable success on the criterion of each member of an incoming group of, say, 200 from the data acquired by administering these three tests, the task would be an onerous one.

Fortunately, however, prediction from a linear regression equation may be facilitated by a very simple charting.

Let us first consider a few facts with regard to our variables by consulting Table 1:

Variable 1. The highest possible score for test 1 is 550. This gives the highest possible weighted value as $550 \times 0.06 = 33$.

Abbreviated original data for the criterion and for three tests given 220 students

TABLE 1

SUBJECT NUMBER	TEST 1 VARIABLE 1	TEST 2 VARIABLE 2	TEST 3 VARIABLE 3	CRITERION 0
1	247	237	52	61
2	247	239	63	56
3	288	234	59	61
4	200	231	68	42
5	260	225	46	69
6	165	203	54	43
.....	.....	.....	.....	.....
220	305	221	51	59
Highest possible score	550	250	75	75
Lowest possible score	0	0	0	0
High score on test	520	239	73	69
Low score on test	124	192	33	36
Range of scores	396	47	40	33
Standard deviations	13.66	14.40	7.43	9.81
Mean	290	218	53	52

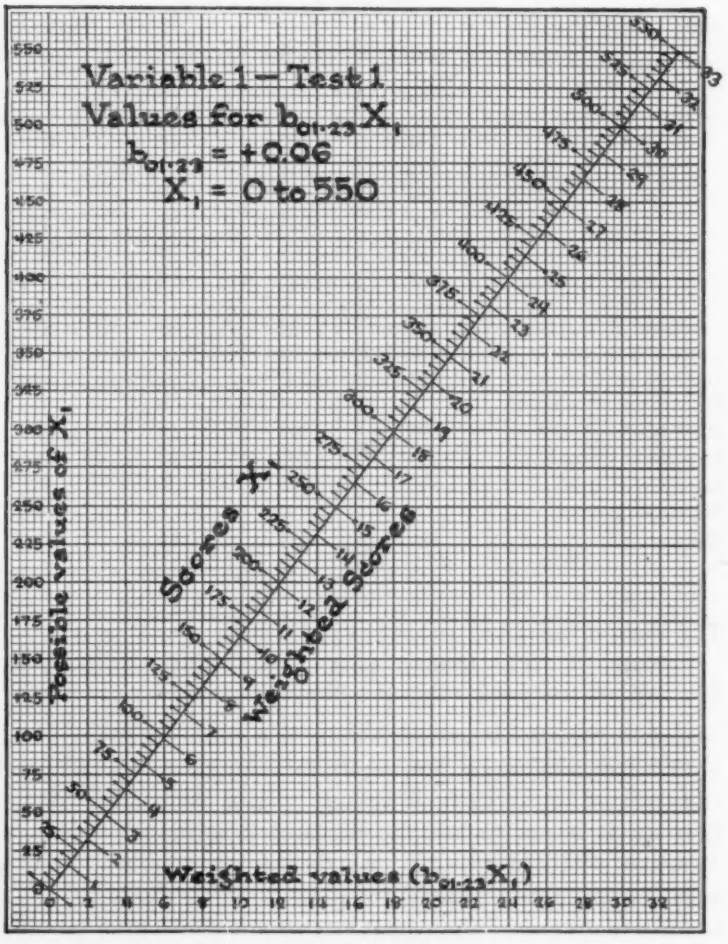
Variable 2. The highest possible score for test 2 is 250. The highest weighted value is thus 107.5.

Variable 3. The highest possible score for test 3 is 75. The highest weighted value is -65.25.

If one wishes, the constant (in this case, -7.35) may be eliminated readily by adding it to one of the independent variables or by dividing it among various variables. Where possible, it is well to use the constant to reduce or eliminate negative weighted values.[1] As the constant here is itself a negative and small, this is not possible. For convenience, it is decided to round the constant to -7.5 and add it to variable 2. This will mean that every weighted value for test 2 will be reduced by 7.5, which makes the highest possible weighted value on test 2 at 100.

[1] Hull, Clark L. *Aptitude testing.* 1st ed. Yonkers: World Book Company, 1928.

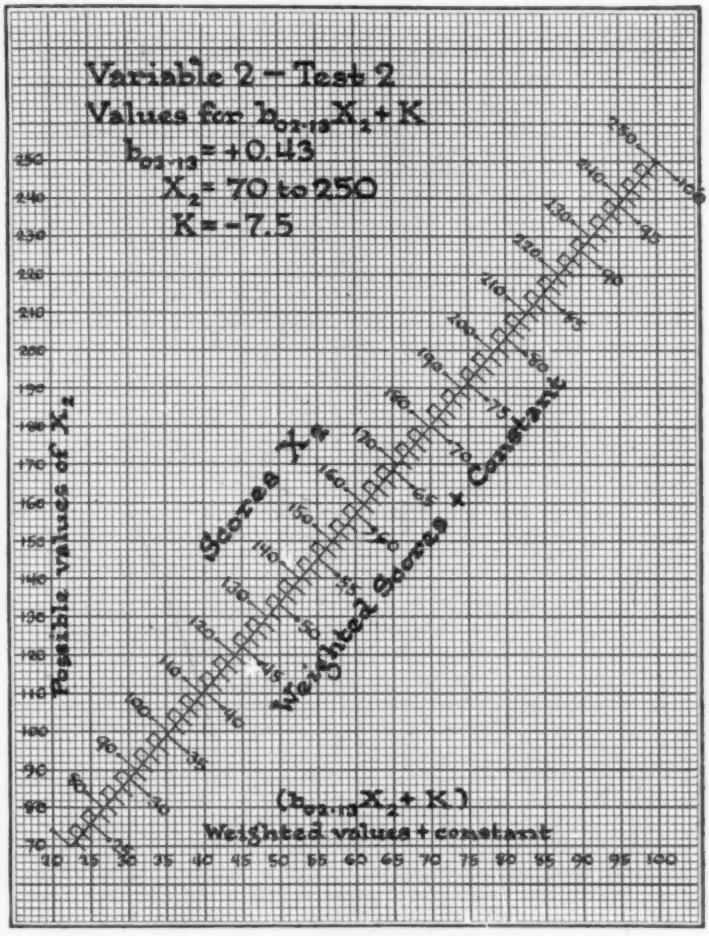
CHART I
Constructing a Prediction Chart (Variable 1)



Now on a sheet of coordinate paper prepare a first quadrant and lay off values for the scores (0 to 550) on the y-axis for variable 1; and for the weighted values (0 to 33) on the x-axis. (See Figure 1.) Locate the point, $550 \times 0.06 = 33$, on Figure 1. Connect this point with the point of origin. Now

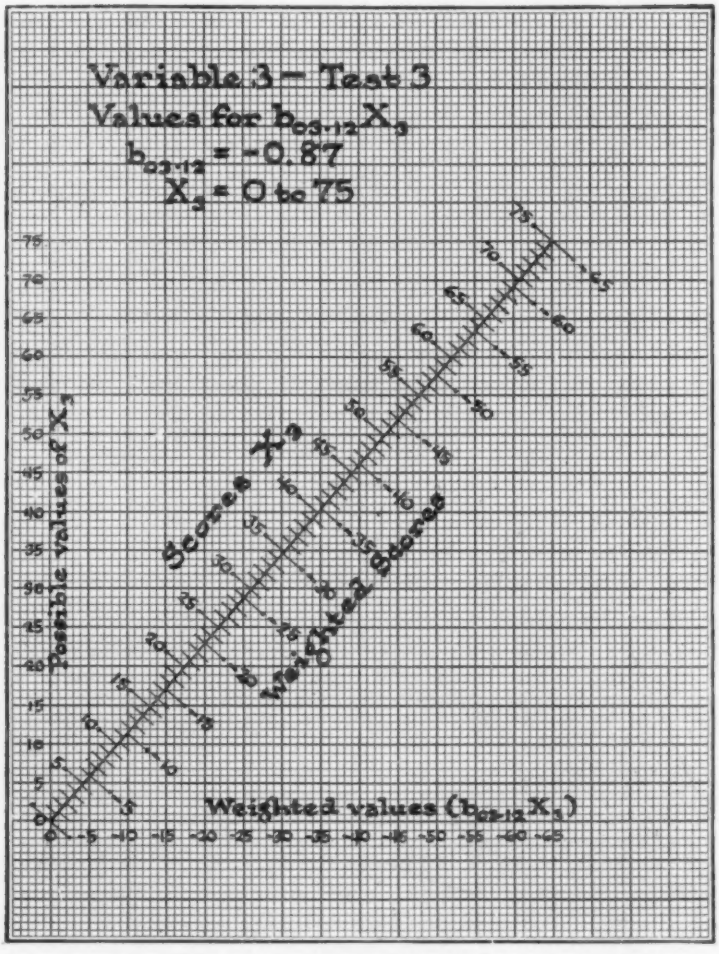
CHART II

Constructing a Prediction Chart (Variable 2)



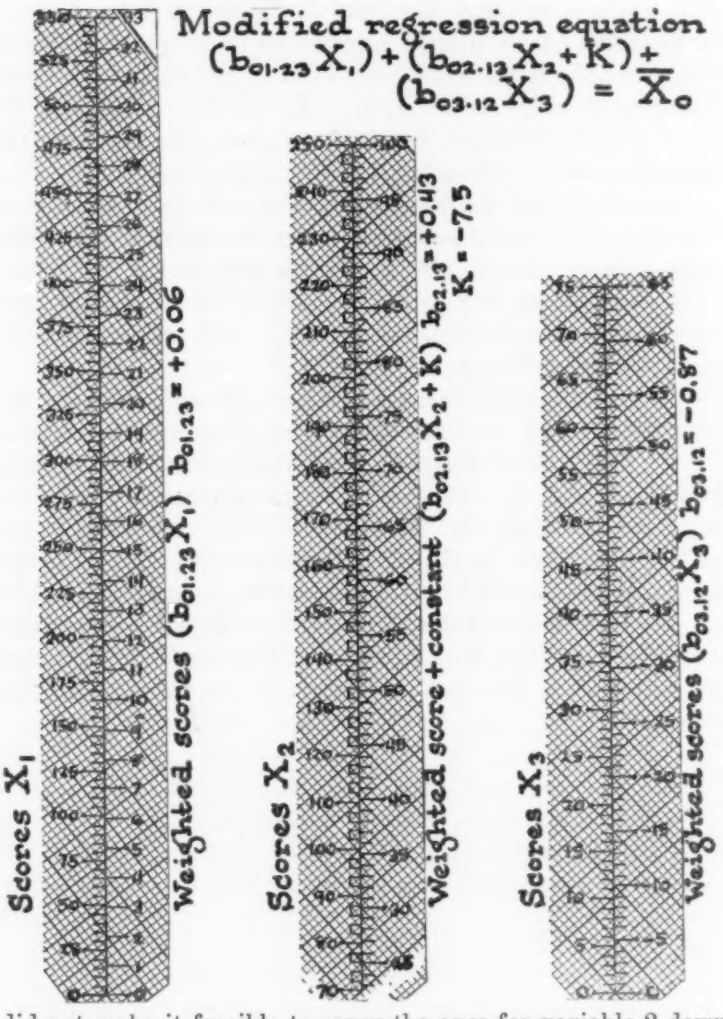
project the values of X_1 on this diagonal axis, forming a scale of scores for X_1. Likewise project the weighted values to the other side of the axis, forming a scale of weighted values. We have now a stationary scale (Chart I) from which values for X_1 may be immediately interpreted into weighted values.

CHART III
Constructing a Prediction Chart (Variable 3)



Scores on variable 2 are charted in a similar manner. Chart II shows the construction of the scale for reading weighted values of X_2 according to our decision to eliminate the constant—$0.43X_2 - 7.5$. The size of the coordinate paper

CHART IV
The Completed Prediction Chart


Modified regression equation
$(b_{01.23}X_1) + (b_{02.13}X_2 + K) + (b_{03.12}X_3) = X_0$
Scores X_1
Weighted scores $(b_{01.23}X_1)$ $b_{01.23} = +0.06$
Scores X_2
Weighted score + constant $(b_{02.13}X_2 + K)$ $b_{02.13} = +0.43$
$K = -7.5$
Scores X_3
Weighted scores $(b_{03.12}X_3)$ $b_{03.12} = -0.87$

did not make it feasible to carry the axes for variable 2 down to the 0-value. However, if the groups we are likely to predict in the future from this data are comparable with the

group we have already measured in establishing these norms, then it is unlikely that there will ever be many scores below 177, which is 3σ below the mean of 218 for variable 2, the σ being 13.66. In a normal distribution a range of 3σ below and above the mean would include 99.73 per cent of the cases. For good measure, however, we shall set the low limit at 70. As $70 \times 0.43 - 7.5 = 22.6$, this will be the starting point for the scale of weighted values for variable 2 on Chart II.

Chart III gives the construction for variable 3, the last of the three facilitating scales necessary in charting the linear regression equation in a four-variable problem.

The scales may then be cut out and mounted on a card for convenience in reference. Chart IV is such an arrangement of the scales on Charts I, II, and III.

We shall illustrate the use of Chart IV by predicting the score of subject 1 on the criterion from his scores on the tests. On the first scale of Chart IV we find the score made by subject 1 on test 1, $X_1 = 247$, and read the approximate weighted value for this score, 15. We follow the same procedure on the scale for variable 2, $X_2 = 237$, therefore $0.43 \times 237 - 7.5 = 94$ by direct reading of the weighted value opposite 237. Also on the scale for variable 3, $X_3 = 52$, with a corresponding weighted value of -45. Adding these three weighted values, we obtain 64 as the most probable score for subject 1 on the criterion.

THE PRIMARY CYLINDER BOARD

EDNA WILLIS McELWEE
New York City Schools

Every psychologist knows that it is often desirable to supplement the Stanford-Binet examination with a non-language test, such as the Seguin Form Board or the Whitmer Cylinder Board. If the examination is given in a clinic such tests are available, but if the psychologist is examining in a school, her equipment is limited to what she can carry with her.

There are not many performance tests small enough to be carried in a brief case, and so supplementary examination is more or less dependent upon the cooperation of the parent in bringing the child to a central clinic. Frequently the examination is never satisfactorily completed.

DESCRIPTION OF THE TEST

The Primary Cylinder Board is nine inches long, four inches wide and one inch high. It is elliptical in shape. The board itself is brown while the cylinders are white with a colored knob. The nine cylinders are all the same depth but of varying diameters.

The largest cylinder is placed in the center and the others are grouped around it in the form of an ellipse. The cylinders are not arranged according to size. The second and third largest cylinders are placed at the right and left of the center. The smallest and next to smallest cylinders are above and below the center. The others are placed between a large and small cylinder.

DIRECTIONS FOR GIVING THE TEST

In giving the test, the Primary Clinder Board is placed near the edge of the table with the red cylinder nearest the child. The cylinders are removed while the child watches and placed

on the table around the upper edge of the board in the following order, beginning at the left of the child: pink, orange, red, yellow, green, dark blue, light blue, lavender and violet. In this way each large cylinder has a smaller cylinder on either side of it and no cylinder is near the opening into which it fits.

The instructions to the child are, "Some of these (pointing to the cylinders) are large and some are small. I wonder if you can find the place where each one just fits. Take them one at a time and put them back for me as quickly as you can." The stop-watch is started as soon as the child picks up the first cylinder.

Nothing further is said unless the child makes an error. Then his attention is called to it by moving the cylinder around in the opening and saying, "This one is too small to go here. Find the place where it fits." If necessary remove the cylinder and hand it to the child.

Sometimes two or three errors are made during the first trial and it may be necessary for the examiner to make the same correction on each occasion. Sufficient help should be given during the first trial so that all cylinders will be placed correctly.

For the second and third trials the cylinders are again removed by the examiner and placed in the same sequence around the upper edge of the board. Before the second trial the examiner says, "Now you know how to play the game. See if you can play it better." Before the third trial, "That was better. Try to put them back a little faster this time."

No help is usually given during the second trial. If, however, no help was given during the first trial, and the child needs help during the second trial, it may be given. Sometimes it happens that the child does not have the experience during the first trial of placing a cylinder in an opening too large for it.

If help was given during the first trial and the child cannot complete the second trial unaided, the examiner removes all

the cylinders for the third trial and says, "Let's begin again. You will do better this time."

SCORING THE TEST

The number of seconds is recorded for each trial. It is best to make a notation on the score sheet in which trial help is given. If no help was given in any of the trials, the best of the three trials is counted. If help was given in the first trial, the better of the second and third trials is counted. If help was given in the second trial instead of the first, the better of the first and third trials is counted. The score in seconds is then changed into mental age by consulting the table of norms.

STANDARDIZATION OF THE TEST

The Primary Cylinder Board was standardized for ages five to nine inclusive. The children to whom the test was given were in the first four grades and kindergarten at Public School 208, Brooklyn, New York, and the kindergarten of Manchester school in Los Angeles, California. There were 75 children in each age group, with the exception of the five year group in which there were 115 children.

In preparing the table of norms the number of cases in each age group was first subdivided into three equal groups according to chronological age. The median score was then found for the first four, the middle four and the last four months of each year. These median scores were then plotted graphically and the scores for the other months were read from the chart. In each case there was such a close relation between the score and the chronological age, that it was unnecessary to smooth any of the curves. The table of norms, see page 416.

RELIABILITY OF THE TEST

The reliability of the Primary Cylinder Board was found by correlating the scores made by 350 children on the second and third trials. The scores of the entire group could not be used because in some cases the number of seconds on one of the trials exceeded 43 or was less than 15.

NORMS FOR THE PRIMARY CYLINDER BOARD

SCORE ON BEST TRIAL	MENTAL AGE	SCORE ON BEST TRIAL	MENTAL AGE
♯45	4–10	29	6–4
♯44	4–11	28	6–6
43	5–0	27	6–8
42	5–1	26	6–10
41	5–1	25	7–0
40	5–2	24	7–2
39	5–3	23	7–4
38	5–4	22	7–6
37	5–4	21	7–10
36	5–5	20	8–2
35	5–6	19	8–6
34	5–7	18	8–10
33	5–9	17	9–2
32	5–10	16	9–6
31	6–0	15	9–10
30	6–2	♯14	10–2

♯ These scores are only approximate, as they were found by extending the growth curve.

The cylinder scores were first evaluated in terms of mental age. Four months was taken as the unit for plotting the correlation. Using the Product-Moment formula, the self correlation was found to be $.78 \pm .014$. By applying the Spearman Brown formula the self correlation was increased to .91, which is a satisfactory measure of reliability.

VALIDITY OF THE TEST

One of the measures of the validity of a test is its relation to chronological age, because mental development in childhood is known to be definitely related to chronological age.

The correlation was found between the best score made on the cylinder test by 400 children and their chronological age. Before plotting the correlation the scores were first evaluated in terms of mental age. Again four months was taken as the unit of measure. Using the Product-Moment formula the correlation between the scores on the cylinder test and chrono-

logical age proved to be .82 ± .01. This is a close relationship, but it is not high enough to preclude the possibility of other factors being measured as well as chronological age.

Another measure of validity is the degree of overlapping of adjacent age groups. The 50th and 75th percentiles of each age group were found and the percent of cases in the next higher age group falling below these two scores were calculated. These results are shown in the following table.

YEAR	50%ILE SCORE IN SECONDS	PERCENT IN NEXT HIGHER AGE GROUP FALLING BELOW	75%ILE SCORE IN SECONDS	PERCENT IN NEXT HIGHER AGE GROUP FALLING BELOW
5	35	6	30	18
6	25	26	24	28
7	22	17	20	24
8	18	26	17	52

A small percent of overlapping at the seventy-fifth percentile level indicates a development with age in whatever factor is measured. A large percent of overlapping indicates little change in the factor measured with developing age. A satisfactory degree of overlapping should not exceed thirty to thirty-five percent.

The fact that 52 percent of the nine year old children fall below the 75th percentile score of the eight year old children indicates that the cylinder test does not discriminate closely beyond the eight year level.

The very small percent of six year old children falling below the 75th percentile score of the five year old children indicates that the cylinder test has great discriminative value at the ages of five and six. In combination with other tests it should be useful as a means of classification of first grade children.

CORRELATION WITH AN INTELLIGENCE TEST

The Haggerty Intelligence Examination Delta I was given to 130 of the eight and nine year old children in the study.

Their IQ's ranged from 70 to 149 with a median IQ of 95. The correlation was found between the mental age on the Haggerty Examination and the best score on the cylinder test made by these children. The cylinder scores were first evaluated in terms of mental age and the unit of four months was used in plotting the correlation. The Product-Moment formula showed a relationship of $.24 \pm .05$, which indicates that the cylinder test measures other factors to a greater degree than it does intelligence.

CORRELATION WITH A PERFORMANCE TEST

The Seguin Form Board was given to fifty of the six-year-old children in the study. The correlation was found between their scores on the Seguin and on the Primary Cylinder Board. Because the norms for the Seguin are in terms of years, that interval was used in plotting the correlation. The scores on both tests were first evaluated in terms of mental age. Using the Product-Moment formula the correlation was found to be $.37 \pm .06$. While this relationship is not large enough to warrant the use of Primary Cylinder Board in place of the Seguin Form Board, yet some common elements exist in the two tests. Both are primarily tests of spatial relationship and visual-motor coordination.

CONCLUSION

The Primary Cylinder Board is a form board small enough to permit its being carried by the psychologist as part of her equipment. It tests visual-motor coordination to a greater degree than it does intelligence. Its self-correlation is .91 and its correlation with chronological age is .82.

It should be useful as a supplementary test along with the Stanford-Binet, especially with younger children when there is a language handicap or auditory or speech defects. It is also interesting and attractive to the uncooperative child. It has unusual discriminative value between ages five and six. In combination with other tests it should prove a reliable means of classifying first grade children.

CONCERNING ONE CRITERION FOR THE CHOICE OF PRIMARY READING TESTS

FLORENCE BAKER
Public Schools, Escondido, California
M. E. BROOM
University of Southern California

Symonds tells us* that the first concern of one who wishes to choose a standardized test is to make sure that the test really measures what it purports to measure. The adequacy and detail with which a test is a measure of a trait, function, or school subject is called its validity. Where several tests have been devised to measure one specific variable it is possible to determine the relative validity of the separate tests by the use of the correlation technique.

The present report is concerned with data yielded by six standardized silent reading tests which were administered to ninety pupils in grade 2B, in the Escondido, California, Public Schools, in December, 1930. The tests were the Los Angeles Primary Word Recognition Test, Form 1; the Detroit Reading Test, Test I, Form A; the DeVault Primary Reading Test, Form 1; the Los Angeles Primary Reading Test, Form 1; the Williams Primary Reading Test, Form A; and the New Stanford Achievement Test, Primary Examination, Reading Section, Form V.

Eighty-seven pupils reacted to the Los Angeles Primary Word Recognition Test. The scores ranged between 1 and 49 points with a mean of 23 and a standard deviation of 11.6. Eighty-seven pupils reacted to the Detroit Reading Test. The scores ranged between 0 and 24 points with a mean of 7 and a standard deviation of 6. Eighty-nine pupils reacted to the DeVault Primary Reading Test. The range included scores from 4 to 43. The mean was 24 and the standard

* Symonds, P. M. "Measurement in Secondary Education." New York: The Macmillan Company, 1927, Page 279.

deviation, 9.7. Ninety pupils reacted to the Los Angeles Primary Reading Test. The range included scores from 0 to 11. The mean was 4 and the standard deviation, 2.6. Eighty-nine pupils reacted to the Williams Primary Reading Test. The scores ranged between 0 and 27 points with a mean of 9 and a standard deviation of 8. Ninety pupils took the New Stanford Primary Examination Reading Test. The range included scores from 6 to 65, with a mean of 29 and a standard deviation of 14. Comparison with end-of-semester 2B norms will show that these pupils were below the grade standard on all tests.

Table I includes coefficients of validity for the several tests. The table follows.

TABLE I.

Measures of relationship between the scores yielded by the several tests

	LOS ANGELES PRIMARY READING TEST	DETROIT	DEVAULT	WILLIAMS	NEW STANFORD
Los Angeles Primary Word Recognition Test	.665	.728	.824	.716	.657
Los Angeles Primary Reading Test		.631	.633	.751	.677
Detroit			.578	.755	.571
DeVault				.781	.599
Williams					.537

These several tests validate each other quite well. Almost any one may be used apparently as an adequate measure of silent reading achievement in grade 2B. More extensive studies are needed before one is warranted in drawing general conclusions. The fact that only one half-grade group reacted, coupled with the retardation in silent achievement in this group which has been noted above, prevents the writers from attempting to read too much significance into the findings given here.

A CRITICAL STUDY OF THE GRAMMATICAL ERRORS OF JUNIOR HIGH SCHOOL PUPILS[1]

GRACE S. RODGERS

John Adams Junior High School, Los Angeles

The problem of this study was two-fold: (1) to determine the extent and nature of the grammatical errors being made by junior high school students of Los Angeles in free composition, and (2) to find a method for effecting a speedy and efficient elimination of such errors.

The first question required a measure of the number of errors. It was therefore necessary to examine a large group of cases taken at random from the varying schools, grades, and IQ levels in a city-wide test.

The second question involved the comparison of identical groups before and after a period of remedial training. There it was not necessary to make a city-wide survey but it was necessary to consider many groups under a large enough number of teachers to overcome the possibility of overshadowing results with personal equations.

In a short study it is not possible to consider more than one remedial method. Hence the method of pre-testing, giving short daily drills, and post-testing was chosen.

PROCEDURE

First Problem. On June 13–15, inclusive, 1928, a test in free composition was given to approximately 29,000 junior high school students of all grades and intelligence levels. The test consisted of a composition—unlimited as to time or length—of a friendly letter in three paragraphs telling about:

First, some interesting event that had recently happened at school.

[1] Prepared under the direction of Dr. A. S. Raubenheimer, University of Southern California.

Second, something that would especially interest the friend to whom the letter was written, and,

Third, the writer's plans for the summer.

These 29,000 papers were separated into groups according to grade. Each grade was then divided into four intelligence levels as follows:

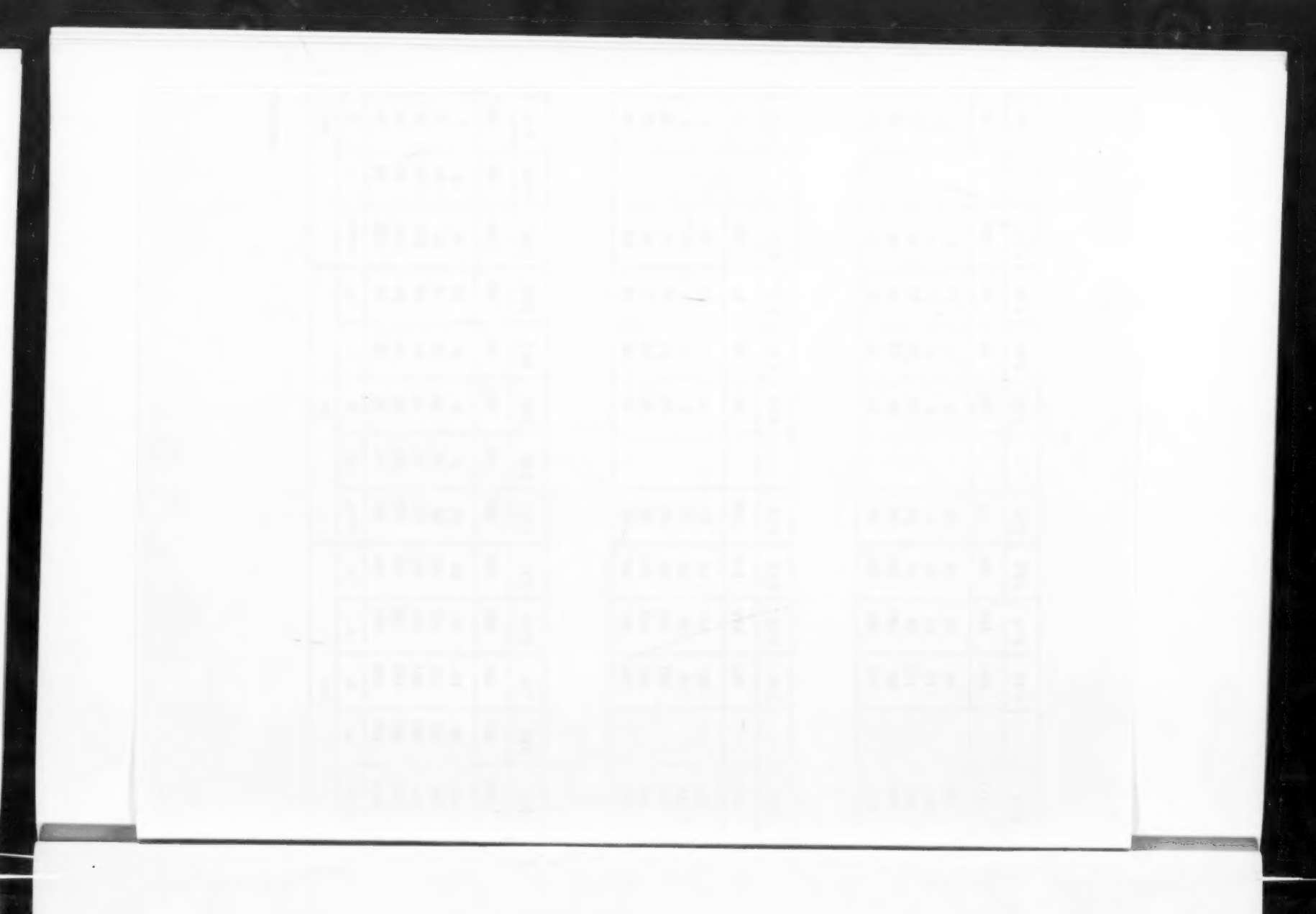
z, below 90
y, 90–110
x, 111–129
w, above 129

Fifty cases were then chosen from each IQ—grade group making 1,200 in all.

Results. The initial tabulation sheets listed 154 different types of grammatical usage for the classification of errors. Forty-eight of these types were early abandoned as negligible. A total of 9,394 errors was recorded under the remaining 106 headings.

Of these 9,394 errors, 7,784, or 82.9 per cent, occurred among only twenty-four types of error. This distribution was as follows:

1.	Capitals with proper nouns and adjectives	1,103
2.	"Will" for "shall" and *vice versa*	1,015
3.	Misuse of verbal contractions	649
4.	Use of needless adverb	527
5.	Omission of preposition	460
6.	Misuse of possessive case	396
7.	Misuse of "to get"	349
8.	Misuse of number	332
9.	Use of terminal preposition	330
10.	Misuse of "to go"	279
11.	Use of "sure" for "surely"	253
12.	Omission of article	226
13.	Failure to use infinitive	241
14.	Misuse of abbreviations	208
15.	Misuse of "lot" and "lots"	199
16.	Use of figures for numbers	185
17.	Omission of the subject	169
18.	Wrong preposition	152
19.	Use of adjective for adverb	147



	B7					A7					
	Z	Y	X	W	Total	Z	Y	X	W	Total	Z
Caps. of Proper N. & Adj.	35	63	42	49	189	38	40	43	44	165	57
"Will" for "Shall"	38	29	43	25	135	37	23	44	33	137	32
Verbal Contractions	13	11	8	13	45	10	3	5	37	55	15
Needless Adverb	24	35	41	16	116	11	33	18	22	84	14
Adjective for Adverb	2	6	10	6	24	4	9	2	3	18	4
Total	112	144	144	109	509	100	108	112	139	459	122
Errors per 1000 wds.	19.9	24.4	20.6	17.8	20.6	17.4	18.5	29.4	17.0	19.2	19.4

Caps. of Proper N. & Adj.	24	33	30		87	24	33	13		70	25
"Will" for "Shall"	8	30	16		54	21	30	22		73	13
Verbal Contractions	4	14	21		39	13	14	14		41	15
Needless Adverb	2	13	3		18	9	13	13		35	5
Adjective for Adverb	9	17	8		34	6	16	15		37	10
Total	47	107	78		232	73	106	77		256	68
Errors per 1000 wds.	18.1	37.9	26.5		27.8	26.2	22.8	20.4		22.8	22.8

Caps. of Proper N. & Adj.	69	22	29		120	21	19	12		52	20
"Will" for "Shall"	20	25	12		57	17	19	16		52	9
Verbal Contractions	8	9	15		32	11	10	15		36	7
Needless Adverb	1	8	3		12	3	5	4		12	1
Adjective for Adverb	15	8	14		37	4	4	6		14	4
Total	113	72	73		258	56	57	53		166	41
Errors per 1000 wds.	30.2	27.0	27.2		28.4	20.8	13.6	13.0		15.2	15.2

TABLE I

mmary of the five principal types of errors made by each IQ-grade group in the diagnostic, pre-, and post-t

Summary of diagnostic results

	B8					A8					B
Total	Z	Y	X	W	Total	Z	Y	X	W	Total	Z	Y	X
165	57	52	46	45	200	56	71	30	41	198	66	58	61
137	32	37	38	41	148	48	41	50	47	186	48	37	54
55	15	23	17	35	90	33	29	43	60	165	37	35	62
84	14	33	13	23	83	11	22	20	33	86	21	31	17
18	4	6	6	7	23	7	9	4	4	24	3	11	4
459	122	151	120	151	544	155	172	147	185	659	175	172	198
19.2	19.4	18.8	17.8	29.1	23.5	23.8	22.5	19.1	22.3	21.8	23.6	20.4	23

TABLE II A

Summary of pre-test results

70	25	26	25		76	24	23	13		60	21	20	20
73	13	32	28		73	33	28	30		91	35	22	28
41	15	19	23		57	7	24	35		66	6	18	20
35	5	7	10		22	7	3	3		13	5	3	8
37	10	11	9		30	8	6	9		23	7	4	4
256	68	95	95		258	79	84	90		253	74	67	80
22.8	22.8	29.3	21.6		24.2	25.4	19.5	19.1		20.9	20.8	15.1	17

TABLE II B

Summary of post-test results

52	20	10	16		46	20	18	23		61	19	16	16
52	9	23	13		45	22	21	22		65	23	22	23
36	7	20	13		40	8	20	42		70	2	8	17
12	1	2	7		10	4	3	2		9	3	2	4
14	4	3	4		11	6	4	7		17	5	2	2
166	41	58	53		152	60	66	96		222	52	50	62
15.2	15.2	21.4	12.0		15.4	17.1	15.7	22.1		18.4	13.3	11.7	14.

and post-tests

B9				A9					ALL				
Y	X	W	Total	Z	Y	X	W	Total	Z	Y	X	W	Total
58	61	34	219	51	31	20	30	132	303	315	242	243	1103
37	54	45	184	53	60	55	57	225	256	227	284	248	1015
35	62	49	183	26	19	31	35	111	134	120	166	229	649
31	17	14	83	20	21	28	16	85	101	175	137	124	537
11	4	6	24	16	8	8	2	34	36	49	34	28	147
172	198	148	693	166	139	142	140	587	830	886	863	872	3451
20.4	23.3	16.1	20.7	19.8	17.9	16.8	15.3	17.4	21.3	20.5	19.1	18.5	19.6
20	20		61	21	23	21		65	139	158	122		419
22	28		85	22	34	26		82	132	176	150		458
18	20		44	14	18	23		55	59	107	136		302
3	8		16	4	6	3		13	32	45	40		117
4	4		15	17	7	7		31	57	61	52		170
67	80		221	78	88	80		246	419	546	500		1466
15.1	17.0		17.4	21.9	18.3	16.6		21.6	22.3	22.5	19.7		21.6
16	16		51	13	17	18		48	162	102	114		378
22	23		68	22	21	13		56	113	131	99		343
8	17		27	13	16	15		44	49	83	117		249
2	4		9	4	2	1		7	16	22	21		59
2	2		9	11	2	6		19	45	23	39		107
50	62		164	63	58	53		174	385	361	390		1136
11.7	14.4		13.1	15.7	19.6	16.7		17.1	18.7	17.2	17.0		17.4

20. Misuse of past tense	135
21. Misplaced adverb	124
22. "&" for "and"	121
23. Misuse of perfect tense	101
24. Misuse of present tense	83
Total	7,784

Another point to be considered was that the greater the number of words written, the greater was the possibility of error. Hence the number of words was counted and the percentage of error for a thousand words figured. The results showed that junior high school students were making approximately twelve errors to a typewritten page of 250 words. It showed no correlation between grade and number of errors, or IQ and number of errors; but it did show that eleven of these twelve errors were made under twenty-four of the 106 types considered.

Second Problem. The second problem was to consider a method for eliminating these errors. For this study twenty-five cases each were used in the three intelligence levels—z, y, x—of each of the six grades. This made eighteen groups of twenty-five identical cases.

The procedure was as follows:

1. A pre-test similar to that given in the city-wide survey was administered, calling for a friendly letter of three paragraphs on:

First, an interesting school event,
Second, Christmas vacation, and
Third, plans for the summer.

2. A drill of three days on each of five types of error was initiated. The errors were:

a. The use of capitals with proper nouns and pronouns.
b. The use of "shall" and "will."
c. The use of verbal contractions.
d. The use of the adverb with special emphasis on the needless adverb.
e. The use of the adjective for adverb, or adverb for adjective.

On the first day a ten minute drill was conducted on the first error listed above.

On the second day a test of ten sentences was given and corrected.

On the third day a review was given on the most difficult phases of the problem under consideration.

Results. The frequency of error per thousand words in the diagnostic, pre-, and post-tests is summarized in the following tables.

TABLE III

Number of words used in the diagnostic, pre-, and post-tests

Number of words in diagnostic test

	Z	Y	X	W	Total	Avg.
B7	5,628	5,899	6,988	6,134	24,649	123.3
A7	5,757	5,844	6,877	8,169	26,647	133.2
B8	6,273	8,052	6,749	6,256	27,330	136.7
A8	6,516	7,641	7,691	8,309	30,157	150.8
B9	7,414	8,420	8,481	9,192	33,507	167.5
A9	8,367	7,769	8,426	9,151	33,713	168.6
Total	38,955	43,625	45,212	47,211	176,003	146.7
Avg.	133.2	145.4	150.7	157.4	146.7	

TABLE IV A

Number of words in pre-test

	Z	Y	X	W	Total	Avg.
B7	2,579	2,821	2,939		8,339	111.2
A7	2,787	4,645	3,776		11,208	149.4
B8	2,989	3,246	4,408		10,643	141.9
A8	3,109	4,305	4,714		12,128	161.7
B9	3,551	4,450	4,710		12,711	169.5
A9	3,759	4,804	4,821		13,384	178.4
Total	18,774	24,271	25,368		68,413	152.0
Avg.	125.2	161.8	169.1		152.0	

TABLE IV B
Number of words in post-test

	Z	Y	X	W	Total	Avg.
B7	3,747	2,672	2,680		9,099	121.3
A7	2,691	4,197	4,063		10,951	146.0
B8	2,695	2,716	4,428		9,839	131.2
A8	3,513	4,212	4,337		12,062	160.8
B9	3,919	4,257	4,298		12,474	166.3
A9	4,021	2,957	3,170		10,148	135.3
Total	20,586	21,011	22,976		64,573	143.5
Avg.	137.2	140.1	153.2		143.5	

They show that a significant decrease has been effected. In three weeks 19.4 per cent or approximately one-fifth of these errors have been eradicated.

Attention is called to (1) the close proximity of the frequency of error in "x," "y," and "z" groups in all three tests; (2) the nearly uniform improvement of all intelligence levels; and (3) the nearly constant improvement of the grade groups in spite of the inherent variation in ability. The most obvious conclusion is that adequate training will effect the desired improvement.

CONCLUSIONS

Summary. The above procedures demonstrated the following facts:

1. In June, 1928, twelve hundred students in writing 176,003 words made 9,394 errors in grammatical usage, or 53.4 mistakes in every thousand words.

2. These 9,394 errors were classifiable under 106 different types of grammatical misuse.

3. Of the 9,394 errors made, 7,784, or 82.9 per cent of them, were classifiable under twenty-four different types of grammatical misuse.

4. Errors of the types classifiable under the twenty-four headings mentioned in 3, occurred with an average frequency of 43.9 per thousand words.

5. In all 3,451, or 37.4 per cent of the errors made were classifiable under five types of grammatical misuse, and these errors occurred with a frequency of 19.6 per thousand words.

6. In January, 1930, 450 students in writing 68,413 words, made 1,466 mistakes in those five types of grammatical misuse mentioned in 5, and these five types of error occurred with a frequency of 21.6 per thousand words.

7. Three weeks later the same 450 students in writing 64,573 words made 1,136 errors in these same five types of misuse; a frequency of error of 17.4 per thousand words.

8. The frequency of error for these five types during a period of three weeks decreased 19.4 per cent.

9. The frequency of error varies within the IQ groups only 2.8 in the pre-test and 1.7 in the post-test; but the average of the IQ groups decreased 3.8 during the period of training.

The conclusions to be drawn from the above facts are at once obvious:

1. Errors are occurring in the written composition of junior high school students with a frequency which merits considerable consideration.

2. All of these errors could be dealt with only over a long period of time and with unprecedented difficulty.

3. The great majority of these errors could be attacked by the intensive study of an extremely small portion of the grammar field.

4. A period of intensive and planned training can materially decrease the frequency of occurrence of these errors.

5. Neither grade nor IQ is a determining factor in the elimination of any particular type of error.

In summary, it seems that a preponderance of evidence has indicated the following conclusion: Although the frequency of grammatical error may be alarmingly great among children of junior high school age, it may be appreciably reduced in all levels of all grades by (1) diagnosing the needs of a class and (2) giving intensive daily drill in the prevalent types of error.

A FURTHER INVESTIGATION OF THE PORTEUS MAZE TEST AS A DISCRIMINATIVE MEASURE IN DELINQUENCY

LOTTA M. KARPELES

Psychological Bureau, State Public Welfare Commission, Providence, R. I.

The present study is an outgrowth of a similar investigation made by Dr. Poull and Miss Montgomery[1] at the New York City Children's Hospital on Randall's Island.

The Randall's Island study, which was made to test the validity of Porteus' claim that scores on the Porteus Maze Test are lower in cases of delinquency,[2] indicated that there is a tendency for socially adjusted children to make slightly higher scores on the Porteus Maze test than they make on the Stanford-Binet, and, conversely, that maladjusted tend to make lower scores on the Mazes than they make on Stanford-Binet.

The authors concluded that the social difficulties of the maladjusted group were due, at least in part, to inability to meet the exigencies of new situations rather than to inability to comprehend or interpret the world about them. Their findings indicated that "anti-social behavior with its concomitant emotional disturbances and defense reactions, may be a result of panic rather than malicious intent."

The subjects in the Randall's Island study were all of the lower levels of intelligence—the range of intelligence quotients being from 42 to 80 in the socially maladjusted group and

[1] The Porteus Maze Test As a Discriminative Measure in Delinquency, JOURNAL OF APPLIED PSYCHOLOGY, Vol. XIII, No. 2, April, 1929, pp. 145–151.

[2] Porteus, S. D., Studies in Mental Deviation, 1923, Training School Bulletin, Vineland, N. J., pp. 75–116.

from 50 to 78 in the socially adjusted group. The average in the former was 63.25 and in the latter 63.

It is the purpose of the present investigation to determine whether or not similar results would be obtained from groups which would more closely approximate a normal distribution, and from groups consisting of subjects on the upper levels of intelligence. A study of various types of delinquency has been undertaken for the purpose of further clarifying the discussion.

The method of procedure[3] in the present study follows that of Poull and Montgomery as closely as possible. However, the intelligence quotient rather than the mental age is used as a basis for comparison. Though the writer found a correlation of – .048 between Chronological Age and Mental Age on the Porteus Maze Test on a previous study with feebleminded subjects,[4] the distribution of Porteus Mental Ages on the present study indicated a positive relationship between these two variables. A statistical check using the product moment method showed the following results:—

Socially Adjusted Group	r = .408
Socially Maladjusted Group	r = .293

The data furnished by the Psychological Bureau,[5] of the State Public Welfare Commission, make possible the compari-

[3] Acknowledgment is due to Dr. Louise E. Poull and Professor H. E. Garrett for consultation on statistical method and to Misses Louise Bauer and Marcia Boynton who assisted in checking the statistical data. An expression of appreciation for assistance in selecting cases is due to Anna I. Griffith, Director of the Children's Bureau, Dr. Joseph H. Ladd, Superintendent of the Exeter School, and Elsie A. Simmonds of the Staff of the State Home and School.

[4] Karpeles, "An Alternative for Stanford-Binet." M.A. thesis, Columbia University, 1927.

[5] The Psychological Bureau examines all individuals committed to the State Home and School for Dependent and Neglected Children pending placement by the Children's Bureau, as well as those who are committed to the institutions for delinquents and to the Exeter School for the Feebleminded.

son of two well defined groups: Conduct disorder cases and socially adjusted cases. The criteria for conduct disorder:—stealing, lying, obstreperous behavior, truancy, fighting and inability to get along with companions or guardians—are those which were used in the previous study. Similarly the control group consists of law abiding, self-controlled, diligent and cooperative individuals. As in the Randall's Island Study, all doubtful cases as well as those in which the scores may have been affected by physical condition or a neurological involvement of any kind have been excluded.

The first and perhaps the most difficult step in the selection of cases was to choose a control group of socially adjusted subjects which would not only fulfill the conduct requirements but which would show also a range of intelligence quotients more nearly approximating a normal distribution than that of the previous study, with an upper chronological age limit of 15 years and 11 months.

In the control group of 185 cases,—110 boys and 75 girls,—the mean chronological age is 11 years 7.8 months. The chronological ages range from 6 years to 15 years and 11 months while the range of mental ages on the Stanford-Binet is from 6 years to 14 years and 11 months, with a mean of 9 years 10.3 months.

The second step was the choice of 185 socially maladjusted subjects. An effort was made to eliminate any great difference in intellectual development by selecting cases, as far as possible, similar in intelligence quotient to those in the control group.

Chronologically, the maladjusted group, in which there are 120 boys and 65 girls, is somewhat older than the control group. The mean chronological age is 13 years 7.3 months, while the range is from 8 to 15 years and 11 months.

In mental age on the Stanford-Binet the maladjusted subjects range from 7 years to 17 years and 6 months. The mean for this group is 11 years 6.2 months.

The cases classified under definite types of offenses, such as

theft, truancy, sex delinquency et cetera were selected at random. Precaution was taken, however, that no case should have any record of any type of delinquency other than that under which it was classified. Therefore none of the individuals in the group of thieves have ever been known to play truant, nor have those classified as truants ever been known to steal. Unfortunately there is some overlapping between the large group representing all types of social maladjustment and the specially classified groups as it was necessary to use the records of some of the subjects under both headings.

The special classifications are as follows:—

The group of 100 thieves included 97 boys and 3 girls.

The 72 truants are all boys.

There are 43 sex offenders, of whom 18 are boys and 25 are girls.

Of the 41 classified as behavior problems, 32 are girls and 9 are boys. The cases in this group include those with records of lying, temper tantrums, enuresis, disobedience, excessive pugnacity and, in fact, all types of conduct disorder not classified under any of the three preceding groups.

To determine whether the upper levels of the intelligence scale would show the same results as the Randall's Island cases, a subdivision of those with intelligence quotients of 80 or over was made in both the control and maladjusted groups. Of the 123 cases in the control group 74 were boys and 49 were girls. In the maladjusted group with intelligence quotients of 80 or over there were also 123 cases of whom 98 were boys and 25 girls.

In each case selected for study records of performance on the Stanford-Binet, Porteus Mazes and the Healy Pictorial Completion Test II were available.

After the selection of the cases the next important step was the computation of the average difference and the reliability of the difference[6] in intelligence quotient for each group on

[6] The statistical procedure was that of finding "The Reliability of the Difference between Two Averages"—See H. E. Garrett, "Statistics in Psychology and Education." New York, Longmans, Green and Company, 1926, pp. 128–133.

the Stanford-Binet and the Porteus Maze Test, as well as to determine statistically, the degree and reliability of the average difference between the various groups in their performance on the Porteus Maze tests.

To accomplish this the average, standard deviation, and the reliability of the average in terms of its standard error were computed for each group, with the following results:—

Socially Adjusted—Entire Group (185 cases)

	Range	Mean	SD	$SD_{(av)}$
Stanford-Binet	45–129	86.2	14.88	1.09
Porteus Mazes	35–189	94.31	23.6	1.13

Socially Adjusted—I.Q. of 80 or over (123 cases)

	Range	Mean	SD	$SD_{(av)}$
Stanford-Binet	80–129	94.45	9.91	.89
Porteus Mazes	45–189	100.55	23.23	2.09

Socially Maladjusted—Entire Group (185 cases)

	Range	Mean	SD	$SD_{(av)}$
Stanford-Binet	50–129	84.62	14.08	1.04
Porteus Mazes	40–139	86.04	17.99	1.32

Socially Maladjusted—I.Q. of 80 or over (123 cases)

	Range	Mean	SD	$SD_{(av)}$
Stanford-Binet	80–129	93.6	9.17	.82
Porteus Mazes	45–139	89.5	16.64	1.50

Socially Maladjusted—Thieves (100 cases)

	Range	Mean	SD	$SD_{(av)}$
Stanford-Binet	50–119	84.4	14.38	1.44
Porteus Mazes	40–129	90.0	16.98	1.70

Socially Maladjusted—Truants (72 cases)

	Range	Mean	SD	$SD_{(av)}$
Stanford-Binet	55–119	83.3	12.42	1.46
Porteus Mazes	45–119	84.5	16.67	1.96

Socially Maladjusted—Sex Offenders (43 cases)

	Range	Mean	SD	$SD_{(av)}$
Stanford-Binet	50–129	80.0	15.71	2.40
Porteus Mazes	45–134	84.6	20.35	3.10

Socially Maladjusted—Behavior Problems (41 cases)

	Range	Mean	SD	$SD_{(av)}$
Stanford-Binet	50–124	78.23	14.08	2.20
Porteus Mazes	50–139	78.60	18.13	2.83

Then, for each group, the average difference, the standard error of the difference, and the reliability of the difference

between the Stanford-Binet and the Porteus Maze Tests were found.

The following tables compare the findings in the Stanford-Binet with those of the Porteus Maze Test, for the groups of 185 cases:

Adjusted	
Stanford-Binet	
Av	86.2
$SD_{(av)}$	1.09
Porteus	
Av	94.31
$SD_{(av)}$	1.73
$SD_{(diff.)}$	2.045
$\frac{D}{SD_{(diff.)}}$	3.966

3.966 ∴ 9999.555 chances in 10,000 that difference is greater than 0

Maladjusted	
Stanford-Binet	
Av	84.62
$SD_{(av)}$	1.04
Porteus	
Av	86.04
$SD_{(av)}$	1.32
$SD_{(diff.)}$	1.68
$\frac{D}{SD_{(diff.)}}$	.845

0.845 ∴ 8023 chances in 10,000 that difference is greater than 0

These tables reveal that there is no significant difference between the scores on the Stanford-Binet and the Maze Test for the maladjusted group, while there is a highly reliable difference in the adjusted subjects.

It would seem, therefore, that adjusted subjects tend to make significantly higher scores on the Mazes than on the Stanford-Binet.

This difference is even more marked in the more intelligent subjects.

The comparison of results with those whose intelligence quotients are 80 and over are tabulated, as follows:—

Adjusted	
Stanford-Binet	
Av	94.45
$SD_{(av)}$	0.89
Porteus	
Av	100.55
$SD_{(av)}$	2.09
$SD_{(diff.)}$	2.272
$\frac{D}{SD_{(diff.)}}$	2.685

2.685 ∴ 9964 chances in 10,000 that difference is greater than 0

Maladjusted	
Stanford-Binet	
Av	93.6
$SD_{(av)}$	0.82
Porteus	
Av	89.5
$SD_{(av)}$	1.50
$SD_{(diff.)}$	1.709
$\frac{D}{SD_{(diff.)}}$	2.399

2.399 ∴ 9918 chances in 10,000 that difference is greater than 0

A study of these tables shows that there is not only a fairly reliable tendency for the more intelligent adjusted children to score higher on the Porteus Mazes than on the Stanford-Binet but also indicates that, conversely, there is an almost equally reliable tendency for maladjusted subjects, on the higher intelligence levels, to score lower on the Maze Test than on the Stanford-Binet.

These findings confirm those of Poull and Montgomery. In the present study the difference in the performance of the adjusted subjects is even more highly significant than in the Randall's Island investigation.

The following tables compare the findings on the Stanford-Binet with those on the Porteus Maze Test for those subjects classified as special types of delinquents.

Thieves

Stanford-Binet	
Av	84.4
$SD_{(av)}$	1.44
Porteus	
Av	90.0
$SD_{(av)}$	1.70
$SD_{(diff.)}$	2.228
$\frac{D}{SD_{(diff.)}}$	2.513

2.513 ∴ 9940 chances in 10,000 that difference is greater than 0

Truants

Stanford-Binet	
Av	83.3
$SD_{(av)}$	1.46
Porteus	
Av	84.5
$SD_{(av)}$	1.96
$SD_{(diff.)}$	2.444
$\frac{D}{SD_{(diff.)}}$	0.491

0.491 ∴ 6879 chances in 10,000 that difference is greater than 0

Sex Delinquents

Stanford-Binet	
Av	80.0
$SD_{(av)}$	2.4
Porteus	
Av	84.6
$SD_{(av)}$	3.10
$SD_{(diff.)}$	3.92
$\frac{D}{SD_{(diff.)}}$	1.173

1.173 ∴ 8790 chances in 10,000 that difference is greater than 0

Behavior Problems

Stanford-Binet	
Av	78.23
$SD_{(av)}$	2.20
Porteus	
Av	78.6
$SD_{(av)}$	2.83
$SD_{(diff.)}$	3.585
$\frac{D}{SD_{(diff.)}}$	0.103

0.103 ∴ 5398 chances in 10,000 that difference is greater than 0

From these tables it will be seen that the thieves tend to make a significantly better score on the Mazes than on

the Stanford-Binet. This is not altogether surprising for stealing does require planning.

In the other three groups of delinquents the difference in performance on the Mazes and the Stanford-Binet is neither significant nor reliable. In addition to this the numbers are too small for final generalization.

As a further check the reliability of the average difference between the Porteus intelligence quotients of the groups was determined.

Comparing the Porteus Average of the entire Adjusted Group with that of the entire maladjusted group we find a difference of 8.27 in favor of the adjusted group. The $SD_{(diff.)}$ was 2.176. A $\frac{D}{SD_{(diff.)}}$ of 3.801 indicates complete reliability.

Comparison of the adjusted group with intelligence quotients of 80 or higher with the maladjusted group of like intellectual status shows a difference of 11.05 with the $SD_{(diff.)}$ of 2.573. The $\frac{D}{SD_{(diff.)}}$ of 4.295 indicates even greater reliability.

Comparing the Porteus average of the entire control group with that of the thieves we find a difference of 4.31 in favor of the former. The $SD_{(diff.)}$ is 2.425. The $\frac{D}{SD_{(diff.)}}$ of 1.777 indicates that there are 9625 chances in 10,000 that the true difference is greater than zero. While this difference is not completely reliable it is significant and shows that socially adjusted individuals tend to score higher on the Mazes than thieves despite the fact that the thieves make higher scores on the Mazes than on the Stanford-Binet.

In comparing the truants' average on the Mazes with that of the socially adjusted group we find a difference of 9.81 in favor of the latter. The $SD_{(diff.)}$ is 2.614 and the $\frac{D}{SD_{(diff.)}}$ of 3.753 indicates that this difference is completely reliable.

Although the next two groups are too small for generalization the results are interesting.

The difference in the Porteus average between sex delinquents and the control group is 9.71 in favor of the latter. The $SD_{(diff.)}$ is 3.55. The $\frac{D}{SD_{(diff.)}}$ is 2.735 and shows that there are 9969 chances in 10,000 that the difference is greater than zero. This is significant though not absolutely reliable.

In comparing the Porteus average of the behavior problems with that of the Socially Adjusted group there is a difference of 15.71 in favor of the control group. The $SD_{(diff.)}$ is 3.317. The $\frac{D}{SD_{(diff.)}}$ of 4.736 indicates complete reliability.

The comparison between the Porteus Average of the special groups of delinquents and that of the entire maladjusted group was not particularly valuable because of overlapping, due to the fact that many cases used in the special groups were also in the latter. It may be interesting to note in passing, however, that the thieves tend to do better on the Mazes than the maladjusted group as a whole, while the behavior problems make a much lower average score on the Mazes than the entire maladjusted group.

SUMMARY

The present study indicates that:—

The control group of 185 socially adjusted children tends to make a significantly higher average on the Mazes than on the Stanford-Binet, while the socially maladjusted group shows no significant difference in performance on these two tests.

The socially adjusted children whose intelligence quotients are 80 and over tend to make a significantly higher average on the Porteus Maze Test than on the Stanford-Binet, while the converse is true of socially maladjusted children on the same intellectual level, who make a lower average on the Porteus Mazes than on the Stanford.

The group of maladjusted subjects classified as thieves achieves a significantly higher average on the Porteus Mazes than on the Stanford-Binet.

There is no significant difference between the averages on the Porteus and on the Stanford-Binet in the group of truants.

Though the groups of sex delinquents and behavior problems are too small for final generalization the results indicate that there is no significant difference in the averages on the Porteus Mazes and the Stanford-Binet with these types of delinquents.

In comparing the average on the Porteus Mazes made by the socially adjusted subjects with that of the maladjusted subjects we find that the adjusted group makes a significantly higher average than the group of delinquents and that this difference is completely reliable.

The difference in favor of the socially adjusted is even more marked and more highly reliable with those subjects on the higher levels of intelligence (intelligence quotients of 80 or over).

The subjects in the control group achieve a higher average on the Mazes than those classified as thieves, truants, sex delinquents or behavior problems. This is true despite the fact that those classified as thieves average higher on the Porteus Maze Test than the maladjusted group as a whole and any of the other classes of delinquents.

CONCLUSIONS

The present study tends to strengthen Porteus' claim that scores on the Maze Test are lower in cases of delinquency and to confirm the findings made by Poull and Montgomery.

It also indicates that on the upper levels of intelligence there is an even more marked tendency for socially adjusted subjects to score higher on the Mazes and for the delinquents to show inferior scores on the Maze Test.

As Poull and Montgomery stated in their conclusions to the Randall's Island Study it seems that "the social difficulty of

these children is not caused by inability to comprehend or interpret the world about them, but is due, at least in part, to inability to meet the exigencies of new situations."

During the years which the writer has spent in working with both delinquent and non-delinquent subjects she has observed that the former not only appear to be less capable of meeting the exigencies of new situations but that they are also less able to profit by their own experiences.

Further study of larger numbers of the various types of delinquents might prove illuminating.

It would be interesting to determine what the results would be with a control group selected from the population at large —non-delinquents who are not State wards.

A similar study should be made to determine whether or not the same results would be obtained with more mature subjects; that is, those who are sixteen years old or over.

On the qualitative side, it might be interesting if not valuable to study the degree in which ability to learn from experience exists with delinquent and non-delinquent subjects.

NOTES AND NEWS

The Child Study Association of America has announced its fall conference for October 17 and 18, 1932, to be held at its headquarters, 221 West 57th Street, and the Hotel Pennsylvania, New York City. Mrs. Sidonie Matsner Gruenberg, Director of the Association, will report on European Educational Conferences of the summer of 1932, including the Sixth World Conference of the New Education Fellowship, at Nice, and the International Conference of Social Work, at Frankfort. Among other speakers will be heard Drs. G. Stoddard, John Dewey and Kimball Young. Specific projects of the Child Study Association will be reported on by various committees.

The first formal summary of the findings of the three-year National Survey of Secondary Education, based on returns from more than 200,000 inquiry forms answered by teachers, pupils, parents and employers, has recently been published. Dr. Leonard V. Koos, of the University of Chicago, noted authority on high school education, directed the Survey.

From August 29 to September 3 will be held in Italy the International Congress of Physiology. It was the desire of some of the members to organize a section on Psychophysiology. Attempts are being made to coordinate the sessions of this Congress with those of the International Congress of Psychology to be held in Copenhagen at almost the same time.

Appointment of Dr. S. J. Beck as Research Assistant in Psychology, Department of Psychiatry, Harvard Medical School, for the year 1932–33 is announced. The appointment will make it possible for Dr. Beck to continue his researches into personality by means of the Rorschach technique. Dr. Beck has resigned as Senior Resident Psychologist at the Boston Psychopathic Hospital, where he has been since 1929.

The World Book Company announces the publication of a new Stenography Test of the series of Blackstone Stenographic Proficiency Tests. Use of the tests will enable the teacher to measure the results in typewriting and stenography classes so that proper readjustment may be made in case of unsatisfactory progress. The tests also make possible a study of the effects of different methods of teaching and of physical conditions.

"Duke University, Durham, N. C., announces a new psychological quarterly to be known as Character and Personality. Publication will begin in September 1932, and continue thereafter in the months of December, March, June and September. A British edition will also be issued simultaneously in London and a German edition in Berlin. Annual subscription price is $2.00. The editor will be Robert Saudek (London) and among members of the editorial board, William McDougall, Alfred Adler, Pierre Janet, C. G. Jung, Johannes Lang."

BOOK REVIEWS

LUMLEY, FREDERICK E. AND BOYD H. BODE. *Ourselves and the World. The Making of an American Citizens.* Illustrations by Guy Brown Wiser. New York, McGraw-Hill Book Co., 1931. viii + 591 pages. $2.00.

With the theses that "democratic society is the kind of society that constantly watches for and continually creates opportunities for individual growth," and "the only worth-while end that human beings know anything about on this earth is *individual growth*," these authors have successfully accomplished the difficult task of writing an unbiased exposition of modern social and political (at least American) organization. Starting with the helpless, asocial baby the writers illustrate the development of the human personality as a social one, and then turn to the social institutions—family, industry, education, press, religion and government—which make our modern world. Throughout the argument rests on modern psychological concepts. In the last section is given a lucid description of our own governments, local, state and national.

The question which kept recurring to the reviewer while reading was "Is this a freshman text, a high school text, or no text at all?" In the hands of a skillful, well informed teacher this book should be an intellectual treat for high school seniors or college freshmen. The content is thought provoking, the presentation is sometimes naïve. The apt and intelligent drawings of Mr. Wiser sound the keynote to each chapter.

C. M. LOUTTIT,
Indiana University.

ALBERT EDWARD WIGGAM. *Sorry But You're Wrong About It.* The Bobbs-Merrill Co., Indianapolis. 1931. 331 pp. $3.00.

Once in a porverbial dog's age a reviewer puts down a book with the exclamation, "I wish I had written that, myself!" I wish I had written this book because it does excellently one of the things that I am firmly convinced very much ought to be done. It aims to convince "Mr. Manstreet"—and this gentleman is not an untutored, eighth-grade graduate with a mental age of twelve, but a host of "lawyers, doctors, preachers, business men, editors, writers, poets, artists, college professors, as well as traveling salesmen, skilled mechanics, and small business men"—that popular beliefs in the fields of psychology and biology are almost uni-

formly wrong, that these beliefs do incalculable damage, and that their prevalence is due to the failure of even seemingly well-trained persons to do their thinking by the analytical methods of science.

Dr. Wiggam sets forth in the early pages four fundamental habits of mind that make it impossible for Mr. Manstreet even to have a right to his own opinions about matters of fact and natural law. He says to him: "You notice the exception to the rule." "You draw general conclusions from particular instances." "You are afraid to let go of your old mental habits for fear if you did you might fall out of the universe." You have a "passion to be important."

In order to drive home to Mr. Manstreet the extent to which these four mental habits lead him astray, the author selects some thirty specific popular notions, and to each he devotes from a paragraph to a score of pages. Some of these notions, such as "that there is an undertow" or "that bald-headedness is due to tight hat bands," are of no special concern to the readers of this magazine (at least professionally!). Others of them, in fact the majority of them, bear more or less directly upon applied psychology. These chapter titles—each following the phrase "you are wrong if you believe"—will illustrate the contents: "That Practice Makes Perfect," "That You Know How to Win an Argument," "That Brilliant Scholars Do Not Succeed in Business," "That Child Prodigies Peter Out," "That Most Great Men Were Born of Poor But Honest Parents," "That You Can Read Human Character at Sight in the Face, Body, and Head"; "That You Can Read Your Character and Destiny in the Stars," "That Telepathy Is a Proved Factor of Nature," "That You Can Improve Your Fundamental Musical Capacities," "That Educating Parents Will Increase the Brain Power of the Children," "That the Beautiful Are Dumb," "That Women Can Drive Automobiles as Well as Men," "That You Can't Reason with a Woman," "That Minister's Sons Usually Go to the Devil," "That One Man Is as Good as Another."

There is no need in this review to enter into the substance of the arguments by which Dr. Wiggam demonstrates the falsity of the beliefs just listed, but it may be pertinent to say that he has taken the pains to verify his assertions by direct reference to men of science and their work, and that he tells his readers who his authorities are and where their work may be found. Further, it may be pertinent to say that he has handled Mr. Manstreet at times in a bluff and breezy manner—witness his proposal (page 79) that the "'literature' (God save the word!") on reading character at sight might be conveniently and descriptively catalogued by the American Library Association as "'The Literature of Damn Fools' . . . even though the introduction to one of these volumes is written by that eminent psychologist and steel manufacturer, Charles M. Schwab." Despite the forthright manner in which

Mr. Manstreet is disemboweled of his pet beliefs, we have an idea he will thank the doctor for performing the operation. And that leads us back to our starting point—that Dr. Wiggam, after all, is more concerned with the causes underlying the patient's condition than with the particular symptoms he may display. Right here is, to my mind, the intrinsic challenge of this book, a challenge that the educational psychologist in especial ought to meet if he can: Is it possible by proper training to imbue, let us say, sixty percent of our high-school graduates, not to mention one hundred percent of our college graduates, with a firmly ingrained attitude toward facts, toward natural law, toward the method and spirit of science such that these graduates will understand the significance of apparent exceptions to the rule, will avoid drawing general conclusions from particular instances, will not be afraid to let go their old mental habits, and will be humble in their search for truth? Is it possible to do this? The question Are we doing it? does not have to be answered; if we were, I would not be led to wish that I had written Wiggam's book.

GUY M. WHIPPLE

HEPNER, HARRY W. *Psychology in Modern Business.* New York, Prentice-Hall. 1930. pp. 728.

We have here the most comprehensive treatment of business psychology thus far published in one volume. It is directed to the business man and the student in the college of commerce and assumes no preliminary training in psychology. This fact should not detract from its interest to the academic psychologist concerned with the applications of the science. He will certainly find in it many points that never occurred to him before and will discover quite a few distinctly original things which the author has done. This originality is typified in the introductory chapter where the author locates the problems of business psychology by obtaining from a goodly number of business men their most pressing problems and then tabulating frequencies of mention.

The book is organized in six parts. Brief mention of some of the topics considered under each will give a rather inadequate notion of the scope. Part I deals with "The nature of the mental life" and begins with evolution and a little neurology. The latter however is somewhat sugar-coated by business terminology, *e.g.*, "the cortex is the control-executive of the brain." In connection with "peculiar people whom we meet in business" some psychoses are discussed and there is likewise stress on the endocrines. Freudian and other mechanisms for avoiding problems are presented and also some method of treating psychopaths, especially by psychoanalysis.

Part II is "Predicting the behavior of the individual." It has a conventional discussion of pseudo-character analysis and a chapter showing the principles making for safety in extending credit. Then follows a discussion of hiring the employee. The interview receives perhaps the greatest stress in this connection although the possibilities of tests, especially intelligence tests, are briefly mentioned. This emphasis is perhaps justified by the fact that most employers use interviews. The author rightly stresses the importance of statistical evaluation of the results. In connection with promotion we have a description of typical rating scales and the ancillary job analyses. The chapter on choosing a vocation is well worth the attention of counsellors. A novel feature is a list of traits for self analysis which includes examples of occupations that utilize the trait in question and also code notations referring to tests for those traits as far as existent. Sample vocational programs worked out for each successive year are presented and discussed.

Part III on "Influencing the behavior of the individual" opens with personal efficiency. The treatment differs from the usual and includes such things as getting rid of salesmen and obtaining promotion in one's concern. Then follow some points about handling employees properly in the interest of morale. The longest chapter in this section deals with salesmanship. It brings out such points as overcoming fear, the approach, demonstrations, sales-talk and has an interesting analysis of different types of prospects with appropriate methods for each. The chapter is a sensible and practical discussion and does not mention the ubiquitous "psychological moment."

Part IV on "Predicting the behavior of the group" lays appropriate stress on market analysis. Methods of testing advertising copy are analyzed and the test-run recommended. It is shown how fashion can be predicted to some extent although it cannot be controlled much through advertising. There is a commendable stress on the need for research in the field of market prediction. Then follows a brief discussion of business cycles and of methods for leveling down seasonal fluctuations in sales.

Part V concerns "Influencing the behavior of the group." Various devices for promoting consumer good-will are discussed and also principles found useful in conducting contests. Then follow two rather extensive chapters on advertising. Much stress is laid on the appeal. The author has a fairly comprehensive classification of appeals that should prove of interest. In the positive-negative controversy he leans somewhat toward the positive appeal. He urges human interest in copy. In connection with actual advertising devices there are sections on coupons, trade names, slogans and their tie-up with the commodity, and

direct mail. Devices for securing attention are mentioned very briefly. The author is evidently more concerned with the ideas involved in the advertisement than with its layout. Three chapters are devoted to supervision of employees. The first is on informal methods and presents the general psychology of the employee, labor propaganda and legitimate methods of presenting the employer's side to labor. Formal methods include welfare work. The author is not especially enthusiastic about it and feels that the employees should be consulted in view of individual differences in sources of satisfaction. Employee representation and profit sharing are discussed. Some good rules for administering suggestion systems are given. In connection with the supervision of women employees the author notes the absence of sex differences in most psychological respects but points out certain problems that confront the executive such as the more temporary character of women's industrial life, their greater concern with personalities, their greater absenteeism and in general the "extra bother." He recommends mixing the sexes in the interest of improving the moral character of the shop.

Part VI deals with psychological research in business. It gives a sound discussion of some of the things to look for when analyzing a research report,—things with which the scientist is perfectly familiar but which probably never occurred to the business man. Then we have some points to be observed in making a questionnaire successful, followed by a briefer discussion of the interview as a research method. There is a final chapter on research "tools" *i.e.*, statistical methods.

The book is an excellent one for its purpose of acquainting the business reader with the whole field, raising problems with which he is familiar and illustrating the solutions with concrete practical examples as far as possible. This "brass tacks" attitude leads to the omission of some fields where psychology has not contributed much of specific value as yet but where interesting beginnings have been made and valuable contributions are very definitely on the way. Morale is a case in point. The reviewer would be inclined to include some such things, point toward the future, speculate a little and get the business reader ready for the contributions when they materialize.

The work does not have, perhaps, the continuity that some readers would prefer. The topics fit logically enough under the five "parts" but there is not a unified treatment with careful transitions. It is rather a collection of a great number of topics with a good treatment of each. The reader can break into the text at most any point and find it understandable and interesting. The desired topic can, of course, be located in the index.

A few special features should be noted: projects or exercises at the end of each chapter many of which are quite original and suggestive;

an occasional cartoon; exercises appealing to the reader's curiosity, with answers in the appendix; epigrams. In the appendix we have a business ability test. A recent choice of the work for the "scientific book of the month" was certainly justified.

MOORE, B. V., AND HARTMANN, G. W. *Readings in Industrial Psychology.* New York, Appleton, 1931, pp. 560.

The present book is a valuable addition to the recent crop of "readings" on various aspects of psychology. It is the first in the industrial field and will be welcomed by all those engaged in teaching this topic. After a brief historical perspective, chapters are devoted to the following: basic principles, popular *versus* scientific procedures, personnel selection techniques, rating scales, mental tests, interests, vocational guidance, training, scientific management, fatigue, working environment, accidents, monotony, morale, motivation, labor unrest, leadership, distribution of product. In each chapter a brief introductory statement sets the stage and is followed by the quotations or selections and then a few suggestive questions at the end. The selections range in length from a single paragraph to ten pages.

No two people would, of course, pick identical selections. The minor variations or shifts of emphasis that occur to the reviewer are not worth mentioning. On the whole the selections are well chosen and the arrangement is logical. The work has greater continuity than one would expect such a compilation to have.

The greatest field for the book will doubtless be as a supplementary reference for courses in industrial psychology where library facilities are limited or the instructor's background is somewhat meager. An experienced teacher might even use it as a text. It has sufficient continuity so that a layman slightly familiar with the field might find it profitable reading. The compilers are to be congratulated on doing a useful piece of work.

ROSSMAN, J. *The Psychology of the Inventor.* Washington, Inventors' Publishing Co., 1931, pp. 252.

The author is a patent examiner and gives us a personnel study of a hitherto neglected vocational group. After an introductory discussion of the rôle of invention in civilization an effort is made to define it. It is distinguished from research in that the latter seeks new knowledge while the former applies existing knowledge.

The book centers largely around a questionnaire sent to patent attorneys, directors of research and individual inventors. Tables are given of the frequency of replies to certain items and there is a judicial use of quotations from those individuals who wrote detailed replies. We

note that inventors possess originality, imagination, perseverance, etc., and that the process of invention involves mental manipulation and trial solution which some inventors get in a flash and others slowly.

The questionnaire itself is not given so it is impossible to evaluate it critically. The reviewer suspects that some of the questions asked may have tended to imply a certain answer. For instance inventors were asked whether their children had shown any tendency toward inventiveness. We would expect a large number of positive replies to such a question. It might be possible to get the desired information in less direct fashion and eliminate the element of suggestion.

The book concludes with a presentation of theories of invention: gestalt, subconscious, self-expression, instinct of workmanship, behaviorism, imitation, inventive faculty.

It is not clear just where the work leaves us. It is not sufficiently systematic and scientific to afford a basis for vocational guidance as to whether a youngster should go in for inventing as a career. It may interest the patent attorney to know something about the persons with whom he is dealing. The inventor may enjoy reading it and comparing his own characteristics with those of other inventors. The book contains little that will be of interest to the professional psychologist.

FORD, A. *A Scientific Approach to Labor Problems.* New York, McGraw-Hill, 1931, pp. 446.

The book is divided into two parts,—first for the reader without technical training and the second including more technical matters, primarily statistical. The order of treatment in the first part is: production measurements, policies of management and selection of employees. In connection with measurement there is a discussion of rating scales and trade tests, pointing out the principal types and the main sources of error. Motion study and job analysis are stressed.

Numerous psychological problems of management are discussed in detail. Among these are remuneration, training for workers and foremen, fatigue, ventilation, noise, accident prevention, illumination, welfare work and maladjustments of psychopathic origin. This section on management is probably the best portion of the book.

Organization of a centralized employment department is outlined. The employment interview receives a considerable stress. A degree of standardization is possible and weighting of the items is desirable. A typical job analysis (carpenter) is given. A warning is very properly sounded against psychological frauds, particularly in the field of character analysis and the use of biographical material for judging character is described. A brief discussion of tests of natural aptitude follows. The selection of tests for discussion in this section is not

especially good. Most of them are in the sensory or motor fields. Only two instances are given of the use of tests in actual vocational selection. The discussion of tests of "character" does not include the more recent objective work.

The second part of the book presents a rather conventional discussion of statistical methods. It includes the use of computing machines and punched card systems. The book concludes with a consideration of present trends in industrial psychology and deplores the overemphasis on tests in the United States.

The author's effort to separate the non-technical from the technical portions of the discussion is theoretically desirable but a very difficult end to achieve. The statistics are eliminated from the first portion to be sure, but other portions, for example that on illumination, get rather technical for the layman. The arrangement likewise makes Part I appear rather thin in spots. For instance there is some discussion about mental tests but nothing about the actual procedure of validating them statistically. Part II does not tie up in any very definite way with Part I. If the technical procedures discussed are not necessary for the continuity of the layman's Part I, the professional reader does not need them anyway because he is already familiar with them.

What seems to the reviewer a major omission is the topic of morale. A number of interesting things on this topic have been published recently. The subject of monotony might well receive a little more space. The author deplores the over-emphasis on tests but in the reviewer's opinion goes too far in the other direction in the light of the fact that the book is quite comprehensive and contains a wealth of other material.

The book is interestingly written and covers a lot of ground. The central portion on management is perhaps the most valuable part of the work. The general reader will find many things of interest throughout. The teacher will find portions worth assignment as collateral reading.

HAROLD E. BURTT,
Ohio State University.

CARROLL C. PRATT. *The Meaning of Music.* McGraw-Hill.

In the search to discover the source of the charm and beauty of music the author first makes an analysis of music. For convenience in study it is divided into three classes—material, form and meaning. The material of music is fundamentally the sensuous substance of tone, and the pleasure derived from listening to beautiful music is the sensuous pleasure of *moving* meaningless sounds given immediately by the very nature of the sensory stimulation produced by acoustic sensitivity. The attributes of tone or pitch, volume, timbre and brightness and it is the distribution of these that constitute the form of music. The characteristic of form is *moving,* or changing, tonal material.

The "Meaning of Music" is a somewhat misleading title for it is found that music has no intrinsic meaning at all. If it comes to have meaning it is through the superficial associations given to it by the listener who has been "contaminated by program-notes and titles." The author maintains that music is not the language of the emotions as is popularly acclaimed (in spite of the description by the publisher that "more than any other art, music is a language of emotion"; emotion is here used in the sense of *moving* musical architecture changing in a given space of time, as contrasted to the im-mobility of painting and sculpture which once perceived do not change). It may or may not produce a given subjective feeling on the part of the listener, depending on how much gratuitous interpretation is given to the musical form. The fallacy, according to the author, of ascribing to music the power of arousing the emotions lies in the confusion of the common characters that exist between auditory and kinesthetic sensation. Thus music (auditory sensation) and emotion (kinesthetic sensation) both possess the dominant characteristic of *movement*,—each may be forceful, weak, restless, calm, indecisive, rhythmic, etc. The perception of such characters in music, being so closely analogous to the bodily accompaniment in emotion, are mistaken for real emotions, which are seldom aroused by music. What music does arouse is the perception of felt *movement* which is the quintessence of, but not necessarily the same as, emotion.

Having denied that the power and attraction of music are connected with the intimate experience of the individual through the emotions, the author proceeds to elaborate Schopenhauer's theory that music is "a copy of the will." Because "ideas of action and *movement* are peculiarly germane to volitional consciousness" and because music likewise is essentially the perception of the *movement* (changing) attributes of tone, it is suggested that music is a copy of the former and for this reason touches life so closely. One wonders if the argument used to refute that music is a language of the emotions would not apply just as well to the theory supported here.

The discussion of the limitations and present status of tests of musical talent, although not connected with the main topic of the book, is nevertheless the best that the reviewer has seen. The book will probably be of more interest to the aesthetic philosopher than to the musician or psychologist.

RALEIGH M. DRAKE,
Wesleyan College.
Macon, Ga.

LANGE-EICHBAUM, WILHELM. *The Problem of Genius* (Translated by Eden and Oedar Paul). New York: The Macmillan Co., 1932. Pp. xix–187. ($3.00).

The author of *Genie, Irrsinn, und Ruhm* (Genius, Insanity, and Fame), published by Ernst Reinhardt, of Munich, in 1928, but not yet trans-

lated into English, has restated recently in brief form the theoretical conclusions which are bolstered in his larger work by encyclopaedic data. *Genie, Irrsinn, und Ruhm* contains the biographies of more than eight hundred individuals who have been regarded as geniuses. In the volume under review, the names of these noted individuals are used for illustrative purposes only, the biographical data being for the most part omitted.

According to Lange-Eichbaum's concept, the problem of genius is a very different problem from that of exceptional and remarkable talent. To Lange-Eichbaum the *talented* individual is one who merely possesses unusual hereditary endowment. Talent therefore is inborn and it may or it may not be recognized. The genius on the other hand is one who brings us values and who is revered by numerous persons. In other words, the genius is, by definition, one who has been recognized as such. Hence, "To speak of 'unrecognized genius' is a contradiction in terms, for genius signifies recognition." p. 68. Moreover, "One cannot be born a genius." p. 24.

Since, according to our author's definition, genius involves recognition, and since recognition is something that is accorded by other individuals, it follows that genius is, in part at least, a sociological phenomenon. To Lange-Eichbaum the genius is not a biologically recognizable variety of human being. " . . . geniuses are created by mankind at large . . . and the making of a genius is a sociological process of becoming. . . ." p. 102. Therefore, "Genius is not an objective entity capable of being studied by the methods applied to the natural sciences." p. 28.

Lange-Eichbaum realizes of course that fame is not always accorded justly; he admits that the halo is not always placed upon the proper brow. For example, "The winning of a battle has not necessarily been due to the strategy of the commander, but perhaps to the skill of those who, months or years before, trained the soldiers. Maybe it was due to rain, or to fog. In like manner, the great statesman is in many cases no more than one factor in a titanic sociological pyramid of forces." p. 16. This being the case, it is not surprising that we find that all possible human types have been raised to the rank of genius, all conceivable varieties and grades of talent, of half-talent, and of no talent at all.

As regards the relationship between nervous instability and genius, Lange-Eichbaum's position is similar to the position that was assumed by William James during the latter's more mature years. This position may perhaps be best described in Lange-Eichbaum's own words. "Grave psychotic disorder, the manifestation of severe psychiatric symptoms, does not, in and by itself, achieve much, if anything at

all.'' p. 125. ''There is not an invariable or necessary association of genius with insanity.'' p. 110. Nevertheless, '' . . . a psychopath is far more likely than a normal individual to be spurred on to self-revelation. The psychopath's feelings are so much more ardent, his impulsive life is so much more mobile, that he is much more likely to get the steam up for a creative task.'' p. 134. ''When a healthy individual sits tranquilly at his desk to compile his memoirs, the result will be very different from what it will be when the demon of a delusion of persecution drives a Rousseau to write the confessions which are to justify him before the world, thus leading to the production of one of the frankest and most straightforward books ever penned.'' p. 135.

Although the author recognizes the weaknesses of the statistical method when applied to the problem of genius, he nonetheless attempts its application. He estimates that among the population at large about 0.5 per cent have grave mental disorders. But among the four hundred individuals that he judged to be our greatest geniuses, he found that from 12 to 13 per cent had been psychotic at least once during their lifetime. Confining his attention to the ''very greatest names,'' numbering seventy-eight in all, he found that more than 37 per cent had been psychotic once during their lifetime; that more than 83 per cent had been markedly psychopathic; that more than 10 per cent had been slightly psychopathic; and that about 6.5 per cent had been healthy.

The proportions of diseased persons became still greater when the author selected the thirty-five persons whom he regarded as ''the greatest geniuses of all''; the psychotics comprised 40 per cent of this group; the psychopathics more than 90 per cent; the healthy 8.5 per cent. On the basis of these findings, the author writes:—''The more profoundly and the more widely an individual is felt by mankind at large to be a genius, the more likely are we to find in this person grave psychopathic symptoms.'' p. 119.

It is apparent at once that both the findings and the conclusions of Lange-Eichbaum are diametrically opposed to those which are set forth by Miss Catherine Cox in Vol. II of the Stanford *Genetic Studies of Genius.* Queries. What is the most plausible explanation for these diametrically opposed findings? Can it be that the two authors selected for their studies different groups of geniuses? (Since the reviewer has not seen Lange-Eichbaum's major study he is unable to answer this question). Is it possible that one or both investigators tended to fasten attention upon, or selected for interpretation, such data as corroborated an *a priori* bias? Is it possible that these conclusions which appear to be as diametrically opposed are in reality reconcilable? And what will be the verdict of history regarding the relationship between nervous instability and genius?

HARVEY C. LEHMAN,
Ohio University.

NEW BOOKS AND PAMPHLETS RECEIVED

Books and pamphlets for review should be sent to James P. Porter, Editor of the JOURNAL OF APPLIED PSYCHOLOGY, Ohio University, Athens, Ohio.

Accidents, Neuroses and Compensation. JAMES H. HUDDLESON. Williams & Wilkins Company, Baltimore, Md., 1932. $4.00. 256 pp.

Biennial Survey of Education in the United States, 1928–30. Vol. II. Office of Education, U. S. Department of the Interior, Washington, D. C. 231 pp.

Criminology. ROBERT H. GAULT. D. C. Heath & Co., Chicago, 1932. $3.48. 461 pp.

Desuggestion. E. TIETJENS. Translated by Eden and Cedar Paul. The Dial Press, Lincoln MacVeagh, New York City, 1932. $4.00. 593 pp.

Economic Theory and Correct Occupational Distribution. HAROLD F. CLARK. Bureau of Publications, Teachers College, Columbia University, New York City, 1931. 176 pp.

Education in the Virgin Islands. ARTHUR E. LINDBORG. Office of Education, U. S. Department of the Interior, Leaflet No. 42.

An Elementary Psychology of the Abnormal. W. B. PILLSBURY. McGraw-Hill Book Company, New York, 1932. $2.50. 375 pp.

Fundamentals of Teaching. G. W. REAGAN. Scott, Foresman & Company, New York, 1932. $2.12. 554 pp.

Managing Minds. CHARLES R. ALLEN and HARRY A. TIEMANN. Century Company, New York. 286 pp.

The Mind in Conflict. RICHARD AMARAL HOWDEN. Oxford University Press, London, England. 2s. 6d. Net. 83 pp.

Muscular Work, Fatigue and Recovery. G. P. CROWDEN. Isaac Putnam & Sons, New York, 1932. $1.00. 74 pp.

The National System of Education in Mexico. CAMERON DUNCAN EBAUGH. Studies in Education No. 16, 1931. Johns Hopkins Press, Baltimore, Md. $2.00. 149 pp.

Number Stories, Book I. J. W. STUDEBAKER, W. C. FINDLEY, F. B. KNIGHT, WILLIAM S. GRAY. Scott, Foresman & Company, New York. 60 cents. 144 pp.

Physical Education and Health Education as a Part of All General Teacher Training Curricula. MARIE M. READY. Office of Education, Bulletin No. 10, 1932. U. S. Department of the Interior, Washington, D. C. 47 pp.

Private Secondary Education in the Association of Colleges and Secondary Schools of the Southern States. ROTHWELL WILCOX. Johns Hopkins University Studies in Education, No. 19, 1932. Johns Hopkins Press, Baltimore, Md. 151 pp.

Psychology in Advertising. ALBERT T. POFFENBERGER. McGraw-Hill Book Company, New York, 1932. 634 pp.

The Psychology and Teaching of Reading. EDWARD WILLIAM DOLCH. Ginn & Company, Columbus, Ohio, 1931. 261 pp.

Report (Third) on Progress in Manchuria in 1932. The South Manchurian Railway, Darien, Manchuria. 235 pp.

Research in Higher Education, Bulletin No. 12, 1931. U. S. Office of Education and University of Oregon. U. S. Department of the Interior, Washington, D. C. 133 pp.

Studies in the Dynamics of Behavior. KARL S. LASHLEY, Editor; CALVIN P. STONE, CHESTER W. DARROW, CARNEY LANDIS, LENA L. HEATH. University of Chicago Press, Chicago, 1932. $5.00. 332 pp.

Directory of American Psychological Periodicals

(Continued)

Psychological Abstracts—Eno Hall, Princeton, N. J. American Psychological Association. Subscription $6.00. 700 pages ann. Edited by W. S. Hunter. Monthly. Abstracts of psychological literature. Founded 1927.

Journal of General Psychology—Worcester, Mass.: Clark University Press. Subscription $14.00 per annum; $7.00 per volume. 1000 pages annually (2 volumes). Edited by Carl Murchison and an international cooperating board. Quarterly. Experimental, theoretical, clinical, and historical psychology. Founded 1927.

Journal of Social Psychology—Worcester, Mass.: Clark University Press. Subscription $7.00. 500 pages annually. Edited by John Dewey, Carl Murchison, and an international cooperating board. Quarterly. Political, racial, and differential psychology. Founded 1929.

Personnel Journal—Baltimore: The Williams & Wilkins Co. Subscription $5.00. 500 pages, approximately. Edited by Walter V. Bingham. Bi-Monthly.

Mental Measurement Monographs—Baltimore: The Williams & Wilkins Company. Subscription $5. 400 pages annually. Edited by Knight Dunlap and Buford Johnson. Published without fixed dates, each number a single research. Founded 1925.

Studies in Psychology and Psychiatry—Baltimore: The Williams & Wilkins Company. Subscription $5. 400 pages annually. Edited by Edward A. Pace. Published without fixed dates, each number a single research. Founded 1926.

Journal of Philosophy—New York: 515 West 116th Street. Subscription $4.00. 728 pages per volume. Founded 1904. Bi-weekly. Edited by F. J. E. Woodbridge, Wendell T. Bush, and H. W. Schneider.

Child Development—Baltimore, Md.: Williams & Wilkins Company. Subscription. $5.00. 400 pages annually. Edited by Buford Johnson. Quarterly. Founded 1930.

Directory of American Psychological Periodicals

American Journal of Psychology—Ithaca, N. Y.: Cornell University. Subscription $6.50. 624 pages annually. Edited by M. F. Washburn, K. M. Dallenbach, M. Bentley, E. G. Boring. Quarterly. General and experimental psychology. Founded 1887.

Pedagogical Seminary and Journal of Genetic Psychology—Worcester, Mass.: Clark University Press. Subscription $14.00 per annum; $7.00 per volume. 1000 pages annually (2 volumes). Edited by Carl Murchison and an international cooperating board. Quarterly. Child behavior, animal behavior and comparative psychology. Founded 1891.

Psychological Review—Princeton, N. J.: Psychological Review Company. Subscription $5.50. 540 pages annually. Bi-monthly. General Psychology. Edited by Howard C. Warren. Founded 1894.

Psychological Monographs—Princeton, N. J.: Psychological Review Company. Subscription $6.00 per vol. 500 pp. Ed. by Raymond Dodge. Published without fixed dates, each issue one or more researches. Founded 1895.

Psychological Index—Princeton, N. J.; Psychological Review Company. Subscription $4.00. 300–400 pp. Edited by W. S. Hunter. An annual bibliography of psychological literature. Founded 1895.

Psychological Bulletin—Princeton, N. J.: Psychological Review Company. Subscription $6.00. 720 pages annually. Psychological literature. Monthly. Edited by Edward S. Robinson. Founded 1904.

Archives of Psychology—New York: Columbia University. Subscription $6. 500 pp. ann. Ed. by R. S. Woodworth. Published without fixed dates, each number a single experimental study. Founded 1906.

Journal of Abnormal and Social Psychology—Eno Hall, Princeton, N. J. American Psychological Association. Subscription $5. Foreign, $5.25. Editor, Henry T. Moore. Quarterly. 448 pages annually. Abnormal and Social. Founded 1906.

Psychological Clinic—Philadelphia: Psychological Clinic Press. Subscription $3.00. 288 pages. Ed. by Lightner Witmer. Without fixed dates (9 Numbers). Orthogenics, psychology, hygiene. Founded 1907.

Psychoanalytic Review—Washington, D. C.: 3617 10th St., N. W. Subscription $6. 500 pages annually. Psychoanalysis. Quarterly. Ed. by W. A. White and S. E. Jelliffe. Founded 1913.

Journal of Experimental Psychology—Princeton, N. J. Psychological Review Company. 500 pages annually. Experimental. Subscription $6.00. Bi-monthly. Edited by S. W. Fernberger. Founded 1916.

Journal of Applied Psychology—Lancaster: The Science Press Printing Co. Subscription $5.00. 400 pages annually. Bi-monthly. Edited by James P. Porter, Ohio University, Athens, Ohio. Founded 1917.

Journal of Comparative Psychology—Baltimore: The Williams & Wilkins Co. Subscription $5.00 per volume of 450 pages. Three volumes every two years. Edited by Knight Dunlap and Robert M. Yerkes. Founded 1921.

Comparative Psychology Monographs—Baltimore: Johns Hopkins University Press. Subscription $5. 400 pages per volume. Knight Dunlap, Managing Editor. Published without fixed dates, each number a single research.

Genetic Psychology Monographs—Worcester, Mass.: Clark University Press. Subscription $14.00 per annum; $7.00 per volume. 1000 pages annually (2 volumes). Edited by Carl Murchison and an international cooperating board. Monthly. Each number one complete research. Child behavior, animal behavior, and comparative psychology. Founded 1925.

(Continued on Third Cover)